المركز الديمقراطي العربي
برلين - ألمانيا

الإحصاء الوصفي والاستدلالي

DESCRIPTIVE AND INFERENTIAL STATISTICS



الأستاذ الدكتور
علي أحمد السقاف
Prof.Dr. Ali Ahmed Alsagf
أستاذ الإحصاء
قسم الإحصاء والمعلوماتية

الطبعة الأولى
2020

رقم التسجيل: VR.3383-6414.B

المركز الديمقراطي العربي
الإحصاء الوصفي والاستدلالي DESCRIPTIVE AND INFERENTIAL STATISTICS

Democratic Arabic Center
Berlin – Germany

هذا الكتاب

هو ثمرة لخبرتي في تدريس مساق الاحصاء الاستدلالي في جامعة عدن، كلية العلوم الادارية قسم الاحصاء والمعلوماتية لعشرة سنوات مضت.

ينقسم علم الاحصاء الى فرعين أساسيين هما: الاحصاء الوصفي والاحصاء الاستدلالي. والهدف من الاحصاء الاستدلالي هو استنتاج خصائص المجتمع من خصائص عينة سحبت منه.

يحتوي الاحصاء الوصفي الاساليب المستخدمة لتلخيص ووصف البيانات الرقمية وذلك بغرض تسهيل تفسيرها. الاحصاء الاستدلالي يحتوي تلك الاساليب والتي من خلالها يتم اتخاذ القرارات حول المجتمع الاحصائي وذلك من واقع العينة المسحوبة من هذا المجتمع.

This book

is the fruit of my experience teaching the so called subject inferential statistics at the University of Aden, Faculty of Administrative Sciences, Department of Statistics and Informatics, for the past ten years.

Statistics is divided into two main branches: Descriptive Statistics and Inferential Statistics. The aim of inferential statistics is to deduce the characteristics of a population from the characteristics of a sample drawn from it.

Descriptive statistics contain the methods used to summarize and describe numerical data in order to facilitate its interpretation. Inferential statistics includes those methods through which decisions are made about the statistical population, based on the reality of the sample drawn from it.





DEMOCRATIC ARABIC CENTER
Germany : Berlin 10315 Gensinger- Str : 112
http://democraticac.de
TEL: 0049-CODE
030-89005468/030-898999419/030-57348845
MOBILTELEFON: 00491742742787 17

الأستاذ الدكتور / علي أحمد السقاف

الإحصاء الوصفي والاستدلالي

Descriptive & Inferential Statistics

الأستاذ الدكتور :

علي أحمد السقاف

أستاذ الاحصاء



الأستاذ الدكتور / علي أحمد السقاف

الناشر

المركز الديمقراطي العربي

للدراسات الاستراتيجية والسياسية والاقتصادية

ألمانيا / برلين

Democratic Arabic Center

Berlin / Germany

المركز الديمقراطي العربي

للدراسات الاستراتيجية والسياسية والاقتصادية ألمانيا/برلين

Berlin10315 Gensingerstr :112

Tel :0049-code Germany

54884375-030

91499898-030

86450098-030

البريد الإلكتروني

book@democraticac.de



رئيس المركز الديمقراطي العربي: أ. عمار شرعان

اسم الكتاب: الإحصاء الوصفي والاستدلالي

الأستاذ الدكتور / علي أحمد السقاف

ضبط وتدقيق: د. سالم بن لباد

التصميم والإخراج: أ. د. بدرالدين شعباني

رقم تسجيل الكتاب: VR . 3383 – 6414 . B

عدد الصفحات: 174

الطبعة الأولى

سبتمبر 2020 م

المحتويات



الفصل الأول

مفاهيم علم الاحصاء
وظائفه وعلاقته بالعلوم الأخرى





وردت كلمة الإحصاء في القرآن الكريم صراحةً من خلال اشتقاقات مختلفة للفعل (حصى) في عدة أمكنة منها : لَّقَدْ أَحْصَاهُمْ وَعَدَّهُمْ عَدًّا (سورة مريم 94) و " إِنَّا نَحْنُ نُحْيِي الْموتى وَنَكْتُبُ مَا قَدَّمُوا وَآثَارَهُمْ وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ فِي إِمَامٍ مُّبِينٍ (سورة ياسين 12) و أيضا ﴿ وَإِنْ تَعُدُّوا نِعْمَةَ اللَّهِ لَا تُحْصُوهَا إِنَّ اللَّهَ لَغَفُورٌ رَحِيمٌ (18النحل)﴾ ثُمَّ بَعَثْنَاهُمْ لِنَعْلَمَ أَيُّ الْحِزْبَيْنِ أَحصى لِمَا لَبِثُوا أَمَدًا ﴿الكهف 12﴾ كما وردت في سورة المجادلة "يَوْمَ يَبْعَثُهُم اللَّهُ جَمِيعًا فَيُنَبِّئُهُم بِمَا عَمِلُوا أَحْصَاهُ اللَّهُ وَنَسُوهُ وَاللَّهُ على كُلِّ شَيْءٍ شَهِيدٌ (سورة المجادلة 6)

في كل الآيات القرآنية المذكورة سلفا ، فإن مفهوم الاحصاء يعني "العد" ولقد استخدم منذ القدم في الاحصاءات التي تهم الحكومات والمؤسسات الادارية المختلفة للدول التي ظهرت في مراحل تاريخية مختلفة . واستخدم الاحصاء في حصر وعد السكان و عدد الجنود والاملاك والانتاج الزراعي والحيوانات وجميع الثروات الاخرى .

أن اصل كلمة إحصاء [Statistics] مكونة من مقطعين هما State وتعني الدولة و stics معناها المتعلقات . و الاحصاء بهذا المعنى يعني متعلقات الدولة . كلمة الاحصاء Statistics] [أتت من الكلمة اللاتينية { Statista } وتعني رجل الدولة الذي يجيد فن الحكم { Statesman}و لقد استخدم هذه الكلمة البروفسور {Gottfried،1719-1772 } الذي عرف الاحصاء بأنه " العلوم السياسية لبلدان متعددة " و كلمة الاحصاء ظهرت لأول مرة في الكتاب المشهور بعنوان

(Baran J.F.Von Bielfed1770) " Elements of Universal Erudition " في احدى فصول الكتاب بعنوان الاحصاء أحتوى تعريفا " بأنه العلم الذي يعلمنا عن ماهية النظام السياسي للدول الحديثة"

الاحصاء هو فرع من الرياضيات التطبيقية . و الطرق الاحصائية وخاصة تلك المتعلقة بالاستدلال عن المجتمع من العينة تبنى على نظرية الرياضيات حول الاحتمالات . و من رواد نظرية الاحتمالات علماء الرياضيات من أمثال James Bernoulli) ، Laplace ، Karl Gauss،) الذين أسهموا في تطوير نظرية الاحتمالات .

العلماء الذين لعبوا دورا كبيراً في تطور علم الإحصاء هم الألماني فر يدريك جاوس (1777 - 1858) ، والفرنسي لابلاس (1749 – 1827) . وللعالم الإنجليزي كارل بيرسون (1857 – 1936) إسهامات كثيرة في علم الإحصاء منها تعريف معامل الارتباط ومعامل الارتباط

الجزئي وتقديره واستخدام اختبار مربع كاي لاختبارات جودة التوفيق والاستقلالية . ويعتبر العالم الإحصائي رونالد فيشر (1890 - 1962) من الذين أضافوا الكثير لعلم الإحصاء ، وهو الذي وضع أساسيات علم تصميم التجارب وتحليل التباين وغيرها من الإسهامات في علم الإحصاء .

1.1 تعريف علم الاحصاء

اكتسب علم الاحصاء أهميته من إمكانية تطبيق نظرياته ، ومبادئه وأساليبه في كل المجالات التي يمكن التعبير عن ظواهره ببيانات يمكن تجميعها . فقد أصبح بالإمكان استخدام الأساليب الإحصائية وتطبيقاتها في مختلف العلوم.

الإحصاء كعلم ونتيجة لعلاقته بعلوم أخرى متعددة يعرف بطرق مختلفة طبقا لاختلاف العلوم ومجالات الدراسة . وظهرت تعريفات متعددة من قبل العديد من الاحصائيين ويمكن ذكر بعض التعريفات على النحو الاتي :

i. الاحصاء هو العلم الذي يتعامل مع جمع ، تصنيف وتبويب الحقائق الرقمية كأساس لتفسير و وصف ومقارنة الظواهر.

ii. علم الاحصاء هو علم التقديرات والاحتمالات

iii. الاحصاء يشير الى الطرق المستخدمة لجمع، وتنظيم وتحليل وتفسير البيانات

بالرغم من الاختلافات الطفيفة في تعريف علم الاحصاء الا ان الاحصائيين يتفقون بكونه "علم جمع وتصنيف وتبويب وتفسير البيانات".

- **الاحصاء الوصفي و الاستدلالي**

I. الاحصاء الوصفي : يحتوي الاحصاء الوصفي الاساليب (الطرق) المستخدمة لتلخيص ووصف البيانات وذلك بغرض تسهيل تفسيرها. وهذه الاساليب يمكن ان تكون بيانيا [graphical] أو التحليل الحسابي و وضعها في جداول مناسبة .

II. خلالها اتخاذ القرارات حول المجتمع الاحصائي وذلك من واقع العينة المسحوبة من هذا المجتمع الاحصائي . وهذه القرارات يتم اتخاذها تحت شروط احتمالية. وتسمى وصف العينة وخصائصها بإحصائية العينة [sample statistics] بينما الخصائص التي تصف المجتمع تسمى معالم . [Parameters]

1.2 مفاهيم علم الاحصاء

1.2.1. العينة والمجتمع الاحصائي

- **العينة** : [Sample] هي جزء من المجتمع ، بحيث تكون ممثلة له تمثيلا جيدا .فمثلا لو تحدثنا عن : طلاب كلية الآداب في جامعة عدن أو اشجار النخيل المثمرة في محافظة حضرموت، أو موظفي البنك المركزي في محافظة عدن كجزء من موظفي القطاع العام في عدن .. فأننا نشير الى العينة . وخصائص العينة تسمى إحصائية [Statistic]
- **المجتمع** : [Population] كلمة مجتمع تستخدم من قبل الاحصائيين للإشارة الى كل العناصر التي تم اختيارها للدراسة . مثلا عندما نقول كل الطلاب في جامعة عدن أو كل اشجار النخيل المثمرة في اليمن أو كل الموظفين التابعين للقطاع العام في محافظة عدن .فأننا نقصد المجتمع الاحصائي . وخصائص المجتمع تسمى معلمة [Parameter]

1.2.2. جمع البيانات

البيانات [Data] هي مجموعة من الحقائق والمشاهدات التي يتم جمعها من مجتمع إحصائي معين . ومن أمثلة البيانات: الاسم والسن والمهنة ومستوى التعليم، ومتوسط الدخل، الحالة الزوجية .....الخ

جمع البيانات يشكل الخطوة الاولى في البحث الاحصائي . و لابد من العناية والدقة في عملية جمع البيانات الاحصائية لا نها تعتبر الاساس في التحليل الاحصائي وذلك لان البيانات الخاطئة تنتج بالضرورة استنتاجات خاطئة و غير موثوق بها .

1.2.3 مصادر البيانات : حسب المصادر فان البيانات الاحصائية تصنف الى بيانات أولية و بيانات ثانوية

البيانات الاولية : بيانات تجمع في الاصل للبحث من قبل الباحث نفسه لغرض الدراسة التي يقوم بها . مثل هذه البيانات اصلية الصفة وتجمع عن طريق المسوحات العديدة التي تجرى من قبل الحكومة وبعض المؤسسات و مراكز الابحاث . البيانات الاولية هي مستقى من المصدر الاول . على سبيل المثال اذا اراد احدهم دراسة الوضع الاقتصادي- الاجتماعي للسكان تحت خط الفقر في قرية معينة ، فأنه يجمع بيانات عن خلفية السكان ، التأهيل العلمي ، دخل الاسرة، عدد المعالين ، الخ ، كل هذه المعلومات تسمى بيانات أولية . لا نه الشخص الاول الذي قام بجمع هذه البيانات . البيانات الاولية

يمكن الحصول عليها من افراد المجتمع كله او جزء منه (عينة) بطريقة مباشرة (ألمقابلة) أو غير مباشرة ، كالبريد والتلفون والانترنت (التواصل الاجتماعي)

• **البيانات الثانوية** : هي البيانات المنشورة او غير المنشورة ، كالكتب والتقارير والمجلات . مثلا الجهاز المركزي للإحصاء ، ومنشوراته تعتبر مصادر ثانوية لمستخدميها ، وكذلك ما يؤخذ من السجل المدني (حالات الوفيات والمواليد، الزواج والطلاق) و كذلك منشورات المنظمات الدولية كالأمم المتحدة.

1.2.4.طرق جمع البيانات : جمع البيانات للأنواع المتعددة من البحوث يمكن اجرائه من خلال الطريقتين التاليتين:

(i) طريقة الحصر الشامل [census method]

(ii) طريقة العينة [sample method]

I. الحصر الشامل : كل مسح يحتوي عملية جمع البيانات المرغوبة من المجتمع الاحصائي والذي يتكون من مفردات(اشخاص ، اشياء ...الخ) . ويسمى المجتمع الكلي تحت الدراسة . اذا كان المطلوب دراسة سكان بلد ما ، يتطلب ذلك فحص واستجواب كل قائمة كل الاسر في مناطق الحضر والريف في ذلك البلد . واذا كان المطلوب دراسة الوضع الاقتصادي للعمال الزراعيين في مديرية ما من محافظة معينة من البلد ،،فأن قائمة المبحوثين تتضمن فقط العمال الزراعيين في تلك المديرية . أن مزايا طريقة الحصر الشامل هي الموثوقية والدقة لان كل مفردة من مفردات المجتمع هنا يتم دراستها . من جانب اخر فأن عيوب الحصر الشامل تمثل في ارتفاع تكاليف الحصر مقارنة بأسلوب العينة وطول الوقت المستخدم.

إن طريقة الحصر الشامل تستخدم في الحالات التالية:

i) اذا كان المجتمع محل الدراسة محدودا

ii) عندما نريد درجة عالية من الدقة

iii) اذا كان موضوع الانفاق على الحصر الشامل والزمن المحدد له لا يشكل معوقا

ii . طريقة العينة : عيوب طريقة الحصر الشامل تم معالجتها في طريقة العينة . وهنا يتم دراسة جزء من المجتمع، ويتم الاخذ بعين الاعتبار بيانات جزء من المجتمع، و بناء عليه فان الاستنتاجات حول المجتمع يتم استنباطها من العينة . والعينة يجب ان تمثل المجتمع تمثيلا جيدا وتعكس صفات المجتمع . و في طريقة العينة ، بدلا من دراسة المفردات المكونة للمجتمع الاحصائي ككل ، يتم دراسة جزء من المفردات تسمى عينة . واذا كانت العينة ممثلة للمجتمع، فان أهم صفات المجتمع يمكن استنتاجها من تحليل العينة .

إن طريقة المعاينة تستخدم في الحالات الاتية:

i) اذا كان المجتمع محل البحث غير محدود
ii) اذا كان المجتمع محل البحث يؤدي الى فنائه
iii) اذا كان الجانب المالي يؤخذ بعين الاعتبار في عملية البحث

- **المميزات الاساسية للعينة**

i. التمثيل : العينة يجب ان تختار بعناية بحيث تكون ممثلة للمجتمع الاحصائي تمثيلا جديا

i. الكفاية : حجم العينة يجب ان يكون كافيا ، خلافا لذلك ، احتمال أنها لا تعكس صفات المجتمع

ii. الاستقلال : كل مفردة من مفردات العينة يجب ان يتم اختيارها بحيث تكون مستقلة عن الاخرى

iii. التجانس : يقصد بالتجانس بأن هناك لا توجد فروق بين المفردات المكونة للمجتمع وكذلك للعينة .اذا اخذت عينتان من نفس المجتمع يجب ان تعطي
نفس النتيجة.

1.2.5. أنواع العينات

تنقسم العينات الى نوعين اساسيين هما:

- **العينة العشوائية [Random Sampling]**
- **العينة غير العشوائية** [Non-random sampling]
- **العينة العشوائية** : تتكون من الانواع الاتية :

i. العينة العشوائية البسيطة [Random Simple Sampling] هي مجموعة جزئية من المجتمع الاحصائي لها نفس الفرصة لتختار كعينة من ذلك المجتمع ، أي بمعنى أن جميع أفراد المجتمع لهم فرصة في أن يُختاروا، ويرجع ذلك إلى أن المجتمع متجانس إذا اختيرت منه عينة وبأي طريقة تستطيع تمثيله وتظهر فيها جميع خصائصه وسماته . و من الجذير بالإشارة هنا بأن كلمة عشوائية لا تعني العفوية ،،بل تعني اعطاء فرص متساوية لكل وحده في المجتمع لان تكون مختارة في العينة .

II. **العينة الطبقية** [Stratified Sampling] العينة الطبقية هي تصميم عينة أكثر كفاءة من العينة التي يتم الحصول عليها عن طريق استخدام أسلوب العينة العشوائية البسيطة ، و يستخدم هذا الاسلوب اذا كانت المجتمعات الاحصائية غير متجانسة في المجتمع الاحصائي الواحد . في العينة الطبقية يتم تقسيم المجتمع الاحصائي الى مجموعة من المجتمعات الصغيرة تميز

بتماثل الوحدات فيما بينها . ثم يقسم حجم العينة المطلوب الى اعداد بنسب حجوم هذه المجتمعات. وكل عدد يمثل حجم عينة عشوائية جزئية توجد في وحدات ذلك المجتمع الصغير ومجموع هذه العينات يمثل العينة المطلوبة

لنفرض أنه مطلوب اختيار عينة مكونة من 200 طالب من مجتمع طلبة كلية العلوم الادارية في جامعة عدن البالغ مجموعهم 5000 طالب . بحيث تمثل جميع الطلبة بصورة متماثلة ، فاذا علمنا ان كلية العلوم الادارية تتكون من اربعة اقسام علمية وهي :

- قسم المحاسبة = 1800 (طالب)
- قسم ادارة الاعمال = 1600
- قسم الاحصاء = 1200
- قسم الادارة الصحية =400
- أذن نصيب طلبة قسم المحاسبة في العينة =(1800/ 5000) x 200 = 72
- نصيب طلبة قسم الادارة في العينة = (1600/ 5000) x 200 = 64
- نصيب طلبة قسم الاحصاء في العينة = (1200/ 5000) x 200 = 48

نصيب طلبة قسم الادارة الصحية = (400/ 5000) x 200 = 16

بعد ذلك نقوم باختيار 72 طالبا بصورة عشوائية من طلبة قسم المحاسبة وذلك بترقيم الطلبة من 1 الى 1800 ثم نختار 72 رقما عشوائيا والمكون من اربعة ارقام عشوائية .وكذلك بنفس الطريقة نقوم باختيار 64 طالبا من قسم الادارة و 48 طالبا من قسم الاحصاء و 16 طالبا من قسم الادارة الصحية .

وهكذا فان حجم العينة المختارة = 72 +64 + 48 +16 = 200

III. **العينة المنتظمة** [Systematic Sampling]

تستخدم طريقة العينة المنتظمة عادة عندما نتوفر قائمة بالمفردات المكونة للمجتمع الاحصائي محل البحث . و يكون اختيار الوحدات منها على أساس تقسيم العدد الكلي للمجتمع على حجم العينة المطلوبة ، ومن ثم توزيع وحدات المجتمع الأصلي وبشكل متساوٍ ومنتظم على الرقم الناتج من ذلك التقسيم . مثلاً: إذا كان العدد الكلي للمجتمع هو (500) طالب وهو رقم يمثل عدد الطلبة في كلية ما، وكانت العينة المطلوبة هي (20) طالبا ، فيكون توزيع الوحدات الكلية الأصلية للمجتمع على الشكل الآتي: (500÷20=25). وعلى هذا الأساس يتحدد رقم العينة

- أي اسم الطالب الأول - بحيث يكون أقل من الرقم(25) وليكن (10) مثلاً .. وتظهر مفردات العينة على النحو الاتي:

10	35	60	85	110	135	160	185	210	235
260	285	310	335	360	385	410	435	460	485

IV. **العينة العنقودية** [Cluster Sampling]

العينة العنقودية تختلف عن المعاينة الطبقية في مبدأ العناقيد . يجب أن تكون العناقيد متباينة في داخلها ، متجانسة فيما بينها أي عكس العينة الطبقية. شكل مشابه للعينة الطبقية. و يتم في هذه الطريقة تقسيم مجتمع البحث إلى مجموعات تسمى عناقيد

[**Clusters**] سواء حسب التوزيع الجغرافي لمجتمع البحث أو بطرق مشابهة. هذه المجموعات تقسم إلى مجموعات إضافية و لهذا السبب أطلق على النوع بالعنقودي بسبب احتواء المجموعات على مجموعات. بعد هذا التقسيم، يقوم الباحث باختيار بعض المجموعات المتحصل عليها بشكل عشوائي، بحيث يتم أخذ جميع أفراد المجموعة المختارة لتصبح جزء في العينة و يتم جمع المعلومات من أفراد هذه المجموعات المختارة.

مثلا اذا اريد التعرف على النمط الاستهلاكي لسكان مدينة عدن فيما اذا كانوا يفضلون الاسماك أم اللحوم .. فانه في المرحلة الاولى يتم الاختيار العشوائي للمديريات والمرحلة الثانية الاختيار العشوائي للمراكز ثم المرحلة الثالثة الاختيار العشوائي للأسر حسب مساكنهم والمرحلة الرابعة الاختيار العشوائي للأفراد من هذه الاسر . وتكون العينة المختارة ممثلة لسكان محافظة عدن باستخدام أسلوب العينة العنقودية .

1.2.6.طرق جمع البيانات الاولية

البيانات الاولية يمكن الحصول عليها وذلك باستخدام الطرق التالية :

i. **المقابلة الشخصية المباشرة** : في هذه الطريقة يقوم الباحثون أو ممثليهم من العدادين بالمقابلة المباشرة للمبحوثين . و تتم عملية أخذ البيانات عن طريق الاستمارة الاحصائية والتي تحتوي على أسئلة محددة تغطي موضوع البحث.

ومن مميزات هذه الطريقة ، دقة البيانات التي يتم الحصول عليها وكفايتها . لكن عيوبها يتمثل في التكاليف العالية والزمن الكبير الذي تتطلبه . ويفترض في العدادين الذين يقومون بعملية استيفاء البيانات ان يكونوا مدربين تدريبا عاليا ولديهم خبرة في عملية المقابلات الى جانب عدم التدخل في الاجابات والتحيز في التأثير على المبحوثين والذي يمكن ان يعطي بيانات مضللة .

ii. **المقابلة الشفوية غير المباشرة** : لا تختلف هذه الطريقة عن المقابلة الشخصية المباشرة من حيث الاتصال المباشر بالمبحوثين للحصول على البيانات. وتكون العملية هنا باستخدام الهاتف كوسيلة اتصال (وسائل التواصل الاجتماعي) بدلا من المقابلة الشخصية . وتكاليف هذه الطريقة اقل بالمقارنة بالطريقة المباشرة و ايضا الوقت المستخدم اقل لأنها تستخدم وسيلة اتصال اسرع .

iii. **طريقة المراسلة:** في طريقة المراسلة يقوم الباحث بأرسال الاستبانة التي تحتوي على الاسئلة المتعلقة بموضوع البحث عن طريق البريد العام و حاليا تستخدم وسائل الاتصال الرقمية كالإنترنت بشكل فعال.

طريقة المراسلة تتطلب درجة عالية من وضوح الاسئلة في الاستبيان وبساطتها وسهولة الاجابة عليها والاسئلة المركبة يتطلب ايضاحها بتعليمات ايضاحية .

من عيوب طريقة المراسلة عدم ضمان مشاركة الجميع في استعادة الاستبيان مكتملا ، وعدم الاهتمام بالموضوع ، و سقوط بعض الاستبانات نتيجة لعدم استكمالها . و تكاليف هذه الطريقة اقل بكثير من طريقة المقابلة الشخصية المباشرة و طريقة المقابلة باستخدام الهاتف .

IV. **المشاهدة الشخصية** : في هذه الطريقة يقوم الباحث بمراقبة الظاهرة وملاحظتها من اجل الحصول على البيانات والمعلومات عن الظاهرة موضوع الدراسة . وتستخدم هذه الطريقة بكثرة في عمليات الانتاج وجودتها والسيطرة النوعية والتي تتطلب مراقبة سير الانتاج وتدوين الملاحظات . وكذلك في مجال الطب حيث يقوم الطبيب بمراقبة حالات المريض بشكل متكرر ودوري وتسجيل الملاحظات حول حالاته ووضعه الصحي .

1.2.6.الاستمارة الاحصائية (الاستبيان) [Questionnaire]

أن فعالية و نجاح استخدام طريقة الاستبيان في جمع البيانات يعتمد على التصميم الجيد للاستمارة الاحصائية . ويجب ان تتوفر الشروط الاتية في الاستمارة الاحصائية (الاستبانة)

I. يجب ان تحتوي الاستبانة رسالة مرفقة تبين هوية الباحث والغرض من الاستبيان وايضا درجة سريتها وعدم افشاء المعلومات لأي جهة ما عدا الغرض الذي صممت لا جله وهي البحث العلمي فقط .

II. عدد الاسئلة في الاستمارة الاحصائية يجب ان تكون في الحد الادنى . العدد المحدد لأسئلة الاستبانة يعتمد على موضوع واهداف البحث ، وفي حالة توسع الاسئلة يجب تقسيمها الى محاور .

III. اسئلة الاستمارة يجب ترتيبها منطقيا ،عدم القفز من محور الى اخر عشوائيا ثم العودة اليه مرة اخرى .

IV.الاسئلة يجب ان تكون قصيرة وسهلة الفهم للمبحوث . الا اذا كان المبحوث متخصصا في موضوع البحث ، خلافا لذلك فأن المصطلحات الفنية والتخصصية يجب تجنبها .

V. الاسئلة الغامضة لا تحبذ في الاستمارة الاحصائية لأنها تؤدي الى اجابات خاطئة .

1.3. وظائف علم الاحصاء

أن أهم وظائف علم الاحصاء يمكن تلخيصه في الاتي :

I. التعبير عن الحقائق بصورة عددية واضحة ودقيقة بدلاً من عرضها ، والتعبير عنها بطريقة إنشائية .

II. تبسيط البيانات الإحصائية بعرضها في جداول أو رسومات بيانية ، وذلك لتسهيل فهمها وتحليلها .

III. يسهل عملية المقارنة بين الظواهر المختلفة ،

IV. يساعد في صياغة و اختبار الفرضيات .

V. يساعد في عملية التنبؤ ببيانات مستقبلية.

VI. استخلاص النتائج واتخاد القرارات المناسبة بدرجة عالية من الدقة

VII. يساعد في صياغة السياسات المناسبة .

1.4. مجالات علم الاحصاء

الاحصاء علم اساسي "كأداة تحليل لكل العلوم" لا يمكن الاستغناء عنه . لا يوجد علم من العلوم لا يستخدم الطرق الاحصائية ، أكانت في مجالات العلوم الاجتماعية أو العلوم الطبيعية المختلفة ، " كالصناعة والتجارة والاقتصاد او الاحياء أو النبات او، علم الفلك والفيزياء والكيمياء والتربية والطب وعلم الاجتماع وعلم النفس وعلم الارصاد .

1.4.1 . الاحصاء والاعمال

النشاطات في مجال الاعمال يمكن حصرها في : الانتاج والمبيعات والمشتريات و المالية ،الطاقم الوظيفي ، والمحاسبة والتسويق وبحوث الانتاج ومراقبة الجودة . بمساعدة الاساليب الاحصائية، فيما يتعلق بمجالات الاعمال السالفة الذكر ، فأن الاساليب الاحصائية تستخدم بكثرة وذلك لوفرة المعلومات الكمية والتي يمكن الحصول عليها و هي ذات الاستخدام الهائل لصياغة السياسات المناسبة . والمعلومات يمكن ان تكون بصورة تقارير او محفوظة في اجهزة الكومبيوتر

أو سجلات أو ملفات . وقدرة المدير تتجلى في استخلاص المعلومات ذات العلاقة بالموضوع من البيانات واستخدامها في اتخاذ القرار . على سبيل المثال فأن بحوث التسويق في المنشآت الكبيرة تستخدم بيانات متعلقة بسلوك المستهلك للمساعدة لا نتاج وتطوير منتجات جديدة . مدير الانتاج ينظر الى بيانات مراقبة الجودة ليقرر متى يقوم بالتعديل في عمليات الانتاج و التصنيع . الجداول والخرائط الاحصائية تستخدم باستمرار من قبل مدراء المبيعات لتوفير الحقائق الرقمية . وبالمثل الاسعار بالنسبة للسلع يمكن تثبيتها ، و الاحصاء يحتل اهمية كبرى في ذلك . أساليب تحليل السلاسل الزمنية والتنبؤات في الاعمال يساعد رجال الاعمال في التنبؤ بأريحية وبدقة عالية تأثير الكم الهائل من المتغيرات .

1.4.2. الاحصاء و الاقتصاد

يهتم الاقتصاد بالإنتاج وتوزيع الثروة وكذلك الاستهلاك والادخار والاستثمار والدخل . البيانات الاحصائية والطرق الاحصائية تحتل أهمية كبرى في فهم المشاكل الاقتصادية وصياغة السياسات الاقتصادية . الطرق الاحصائية ، الى جانب انها تساعد في صياغة السياسات المناسبة ، فإنها أيضا تقيم تأثيراتها. على سبيل المثال " ماذا ننتج وكيف ولمن ننتج " هذه الاسئلة تحتاج كما هائلا من البيانات الاحصائية . إحصائيات الانتاج تساعدنا في تعديل الطلب المقابل للعرض . واحصائيات الاستهلاك تمكننا من ايجاد الطريقة التي يمكن من خلالها مختلف طبقات السكان من انفاق مداخيلهم .

ومن العلوم الحديثة نسبيا التي لها علاقة وطيدة بالإحصاء واثرت تأثيرا ملحوظا في علم الاقتصاد هو علم الاقتصاد القياسي[Econometrics] الذي يتضمن تطبيق الطرق الاحصائية في النظرية الاقتصادية ويستخدم بشكل واسع في البحوث الاقتصادية .

1.4.3. الاحصاء والرياضيات

توجد علاقة ارتباط جوهرية بين الاحصاء والرياضيات . التطور الملحوظ في الطرق الاحصائية هو نتيجة لاستخدام مختلف فروع الرياضيات ، كالجبر والتفاضل والتكامل .أن نظرية الاحتمالات ترتبط ارتباطا جوهريا بالتحليل الاحصائي ولا يمكن مناقشة الاحصاء دون الاستيعاب الجيد لنظرية الاحتمالات .

1.4.4. الاحصاء والعلوم الفيزيائية و الطبيعية

العلوم الفيزيائية ، كعلم الفلك ، الجيولوجيا ، والفيزياء ، هي من المجالات التي استخدمت الطرق الاحصائية في فترات مبكرة واسهمت في تطوير هذه العلوم . وتستخدم الاساليب الاحصائية بكثافة في الوقت الراهن في الكيمياء والهندسة ، وعلم الارصاد .

الاساليب الاحصائية استخدمت في مجالات العلوم الطبيعية . فعلوم الاحياء والطب و الحيوان والنبات تستخدم الطرق الاحصائية المختلفة بشكل مكثف مقارنة بالعلوم الفزيائية . فمثلا في الطب لكي يشخص الطبيب المرض يجب ان تتوفر لديه بيانات عن حالة المريض والمتعلقة بضغط الدم و دقات القلب ودرجة الحرارة . وايضا للحكم على مدى فعالية عقار معين لمعالجة المريض فان التجارب يجب ان تجرى للتعرف على نجاح التجربة او فشلها وهذا يتطلب معرفة عدد الاشخاص الذين تعالجوا بعد تناولهم للعقار . في علم الاحياء يجب ان يعتمد على الاحصائيات لا جراء التجارب وذلك للتعرف على مدى تأثر النبات بدرجات الحرارة و التربة .

1.4.5. الاحصاء و البحث العلمي

علم الاحصاء اساسي ولا غنى عنه في البحوث العلمية . والتطور الهائل في المعرفة هي نتاج للتجارب التي اجريت باستخدام الاساليب الاحصائية . فمثلا التجارب التي اجريت حول انتاج المحاصيل الزراعية والانواع المختلفة من الاسمدة و التربة أو تربية الحيوانات طبقا لاختلاف انواع التغذية والبيئة، دائما ما يتم تصميمها واجراء التحليل وذلك باستخدام الطرق الاحصائية . والاساليب الاحصائية تستخدم بشدة في العلوم الطبية واثرت تأثيرا جليا في الطب والصحة العامة.

أن الاستخدام المكثف للأساليب الاحصائية في العلوم الطبيعية والاجتماعية أدى بروز فروع للعلوم الاحصائية تشكل مجالات فرعية منفصلة . الى جانب علم الاقتصاد القياسي الذي يدرس النظريات الاقتصادية وتطبيقاتها كميا و يستخدم ايضا في اثباتها فان هناك علوم فرعية اخرى برزت كالإحصاء الرياضي ، والاحصاء الاقتصادي ، الاحصاء التربوي ، الاحصاء الاجتماعي، والاحصاء الحيوي والاحصاء السكاني والاحصاء الجنائي . كل هذه الفروع تستخدم الاساليب الاحصائية وذلك بالارتباط بالمواضيع التي يتم دراستها في كل مجال من مجالات العلوم الطبيعية والاجتماعية .



الفصل الثاني

تصنيف البيانات الاحصائية وعرضها





الأستاذ الدكتور / علي أحمد السقاف

في الفصل الاول تم تعريف علم الاحصاء ، كما ورد في أدبيات الكثير من الاحصائيين بأنه " العلم الذي يتعامل مع "جمع ، تصنيف و تبويب ،وعرض البيانات الاحصائية وتحليلها وتفسيرها " . وتم ايضا توضيح طرق جمع البيانات الاحصائية ومصادرها . وفي الفصل الثاني سيتم استعراض طرق تصنيف وتبويب وعرض البيانات الاخصائية.

2.1. تصنيف البيانات

2.1.1: مراجعة البيانات : ما أن تنتهي عملية جمع البيانات، حتى تأتي مرحلة المراجعة، والتي تفحص فيها الاستمارات الإحصائية، وتدقق على أن تستبقى الاستمارات ذات الإجابة الصحيحة الكاملة، وتستبعد الاستمارات الناقصة، أو الاستمارات ذات الإجابات غير الصحيحة .

2.1.2 : جدولة البيانات : ومن اجل توفير الجهد والزمن عبر مراجعة البيانات ومعرفة مدى فائدتها للباحث ، وتوافقها مع أهداف البحث ، ثم منها عملية جدولتها . أي وضعها بأصغر حيّز ممكن ، و الترتيب الذي توضع فيه البيانات بعد فرزها وتنظيمها، يسمى بـ (الجدول). والجداول هنا تكون على أشكال مختلفة ومتنوعة، إذ منها الجداول الأولية، ومنها الثانوية . وكل منها يصلح للاستخدام في حالات معينة إلاّ أنها جميعاً تهدف إلى إبراز البيانات وتوضيحها في حجم مكثف ومصغر.

2.1.3: تصنيف البيانات : يعد تصنيف البيانات من الخطوات المهمة في عملية التبويب فبعد الانتهاء من عملية جمع الاستمارات المعنية بالبيانات المطلوبة ومراجعتها تأتي عملية فرز البيانات إلى مجاميع وأصناف صغيرة، توحدها قاعدة معينة كأن تشترك كل مجموعة في بعض من الصفات أو الخصائص (الطول ، العمر ، الوزن ، الحالة الأسرية ، فئة الدخل ...الخ) بحسب متطلبات البحث .

2.1.4 : تبويب البيانات : بعد اتمام تصنيف البيانات نبدأ بعملية تبويب البيانات ويقصد بتبويب البيانات ، عملية تصنيف وتفريغ البيانات في جداول. وللتبويب أساليب مختلفة، يأتي اختلافها بحسب طبيعة البيانات المراد تبويبها، وكذا الكيفية التي سوف تستخدم بها البيانات بعد تبويبها، ومن أنواع التبويب :

- **التبويب الجغرافي** : عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة جغرافية معينة او تقسيم اداري معين.
- **التبويب الزمني** :عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها يعود الى وحدة زمنية معينة كاليوم او الاسبوع او الشهر او السنة.

- **التبويب النوعي** : عبارة عن عملية تجميع البينات وترتيبها في جداول خاصة على اساس ان كل جمع منها يشترك بصفة معينة كالجنس ،الحالة الاجتماعية ،الحالة التعليمية ، الوظيفة . تجدول البيانات في هذا النوع من التبويب بحسب صفة النوع ، والنوع هنا يعبّر عن ظاهرة والصفقة محل الدراسة هنا تتميز بأنها غير قابلة للقياس .

مثال(2.1) ضع البيانات النوعية(الوصفية) الاتية في جدول تكراري مناسب ، والبيانات تمثل عينة من درجات 50 طالبا في مساق الرياضيات لطلاب الثانوية العامة للعام الدراسي 2001/2000.

جيد	جيد	جيد جدا	جيد جدا	ممتاز	مقبول	جيد	راسب	راسب	جيد جدا
مقبول	راسب	ممتاز	ممتاز	جيد جدا	جيد	جيد	جيد	جيد	مقبول
جيد جدا	مقبول	جيد	ممتاز	جيد	جيد	ممتاز	راسب	راسب	ممتاز
راسب	مقبول	جيد جدا	جيد	جيد جدا	جيد	راسب	جيد	جيد	جيد جدا
مقبول	مقبول	جيد	جيد	راسب	راسب	جيد	جيد	مقبول	جيد جدا

كما اشرنا فيما سبق ، بأن الجدول الاحصائي هو عملية تبويب للبيانات الكمية و(النوعية) وتلخيصها عموديا وافقيا بقصد تسهيل دراستها والاستفادة منها . و عند أعداد الجدول الاحصائي ، لابد من مراعاة القواعد الاتية:

I. أن يكون للجدول عنوانا واضحا يعكس محتوياته
II. لابد من ترتيب الاعمدة والصفوف وترقيمها
III. المعلومات الواردة في الجدول تكون حسب مصدر معين يجب ذكره

وبالعودة الى المثال رقم (2.1) ، فان البيانات هي وصفية (نوعية) والتي تمثل التحصيل العلمي للطلاب في مساق الرياضيات . ونلاحظ أن البيانات الواردة في المثال هي بيانات خام ، يصعب الاستفادة منها وعليه لابد من وضعها في جدول يسمى جدول التوزيع التكراري للبيانات الوصفية (أو النوعية) ويتم تكوين الجدول كالاتي :

جدول (2.1) جدول التوزيع التكراري للبيانات الوصفة

الصفات (تقديرات الطلاب)	علامات التفريغ	التكرارات (عدد الطلاب)
ممتاز	//// /	6
جيد جدا	//// ////	9
جيد	//// //// //// ///	18
مقبول		8
راسب		9
المجموع	-	50

- **تبويب البيانات الكمية**

تبويب البيانات الكمية يشير الى تلك البيانات التي يمكن قياسها . ويتم تبويبها طبقا للصفات . مثل الوزن والطول ، الدخل ، المبيعات ، الارباح ، الانتاج ..الخ

البيانات الكمية يتم عرضها في عدة جداول وهي :

i. الجدول التكراري البسيط

ii. الجدول التكراري ذو فئات

iii. جدول تكراري متجمع صاعد

iv. جدول تكراري متجمع هابط

مثال (2.2) البيانات الاتية تمثل ، عدد أطفال خمسين أسرة حسب استمارة الاستبيان من واقع التعداد العام لبلد ما . المطلوب عرض هذه البيانات في جدول تكراري مناسب

3	2	2	1	5	4	2	1	3	4
3	0	2	1	2	3	3	2	1	1
2	3	0	3	2	1	4	3	5	5
4	3	6	5	4	3	1	0	6	5
5	4	3	1	2	0	1	2	3	4

الحل : نفرغ البيانات الخام اعلاه حسب الجدول التكراري أدناه

جدول(2.2) الجدول التكراري لبيانات 50 أسرة

عدد افراد الاسرة	علامات التفريغ	التكرار	التكرار النسبي
0	////	4	0.08
1	//// ////	9	0.18
2	//// ////	10	0.20
3	//// //// //	12	0.24
4	//// //	7	0.14
5	//// /	6	0.12
6	//	2	0.04
المجموع	-	50	1.00

- **التكرار النسبي** : يتطلب أحيانا معرفة نسبة جزء معين من الكل ،للموضوع محل الدراسة،و هذا يسهل عملية التحليل . ويحسب التكرار النسبي بقسمة التكرارات على حجم العينة .

التكرار النسبي = التكرار / حجم العينة

وفي المثال رقم 2.2 تم حساب التكرار النسبي في الجدول لعدد الاطفال في الاستبيان المكون من 50 أسرة

- **الجدول التكراري ذو فئات**

مثال (2.3) تبين القيم التالية الدرجات النهائية لخمسين طالبا جامعيا في امتحان مساق الاحصاء لطلاب المستوى الاول . المطلوب توزيع هذه الدرجات في جدول تكراري مناسب .

74	50	76	62	78	88	57	73	80	65
79	67	65	78	76	65	85	71	75	62
75	95	60	79	83	71	61	89	78	96
57	68	74	69	77	94	72	82	78	66
74	55	87	95	78	75	63	98	72	61

الحل

من أجل تفريغ البيانات في المثال رقم (2.3) في جدول تكراري مناسب فأننا نستخدم الجدول التكراري ذو فئات

ويتم تصميم هذا الجدول طبقا للخطوات الاتية:

I. أيجاد المدى (Range) للبيانات الخام وذلك بطرح أدنى قيمة في البيانات من أعلى قيمة.

المدى = أعلى قيمة – أدنى قيمة

II . نوجد طول الفئة وذلك بقسمة المدى على عدد الفئات

طول الفئة = المدى / على عدد الفئات

ولإيجاد طول الفئة ، لابد ان نحدد أولا عدد الفئات . وتحديد عدد الفئات يعتمد أساسا على خبرة الباحث و عدد المفردات . وعادة تستخدم الفئات بحيث تتراوح بين 5 و 20 ، أي لا تقل عن 5 و لا تزيد عن 20 .

و هناك قاعدة يمكن استخدامها لإيجاد عدد الفئات ، وهي قاعدة سترجس(Struges Rule)

$$K = 1 + 3.3 \log N$$

حيث أن:

k =عدد الفئات

N = عدد المفردات

Log = لوغاريتم العدد

من ثم نوجد طول الفئة (I) طبقا للقانون ادناه : حيث أن L = الحد الاعلى للفئة ، S = الحد الأدنى للفئة ، R = المدى .

$I = L - S / k$

$R = L - S$

$I = R / k$

$I = R / 1+3.322 \log N$

III. بعد حساب طول الفئة فأننا نوجد الفئة الاولى وذلك بإضافة طول الفئة الى أدنى قيمة في البيانات

نوجد الفئات الاخرى بشكل تتابعي الى أن نصل الى أعلى قيمة في البيانات الخام .

بالعودة الى المثال رقم (3) :

I . نوجد المدى (R)

$R = L - S$

$= 98 - 50 = 48$

Ii . نوجد طول الفئة من العلاقة :

$I = R / k$

$K = 1 + 3.322 \log N$

$K = 1 + 3.322 \log (50)$

$K = 1 + 3.322 (1.6989)$

$K = 1 + 5.6437 = 6.6437 \sim 7$

$I = /k = 44/7 = 6.2 = 6$

لكن في مثالنا هذا نفترض ان عدد الفئات (K) =5

$I = 48/5 = 9.6 \quad (k = 5)$

$I = 10$

Iii . الحد الأدنى للفئة = 50

طول الفئة = 10

أذن الفئة الاولى = 50 + 10 = 60

ونحصل على الفئة الاولى = (50-60)

ونضيف طول الفئة الى الاحد الاعلى للفئة الاولى ونحصل على الفئة الثانية

= 60 +10 = 70

أذن الفئة الثانية = (60-70)

وتظهر الفئات بالصيغة التالية

50-60 ، 60 -70 ، 70 -80 ، 80 -90 ، ...... وهكذا

و يظهر شكل جدول التوزيع التكراري ذي الفئات بالشكل الاتي:

جدول (2،3) جدول التوزيع التكراري ذي الفئات لدرجات 50 طالبا في مساق الاحصاء

الفئات	علامات التفريغ	التكرار
50 – 60	////	4
60 - 70	///// ///// ///	13
70 -80	///// ///// ///// ///// /	21
80 - 90	///// //	7
90 - 100	/////	5
المجموع		50

- **التكرار النسبي** [Relative Frequency] و التكرار التراكمي Cumulative [frequency]
  - التكرار النسبي لتقدير يساوي تكراره مقسوماً على مجموع التكرارات.
  - التكرار التراكمي لتقدير يساوي عدد عناصر العينة الذين لهم ذلك التقدير على الأكثر
  - التكرار التراكمي النسبي لتقدير يساوي تكراره التراكمي مقسوماً على مجموع التكرارات

وبالعودة الى المثال رقم (3) يمكن احتساب التكرارات النسبية والتراكمية كالاتي

جدول(2،4) التكرارات النسبية والتراكمية

الفئات	التكرار (f)	التكرار النسبي	التكرر النسبي المئوي (%)	التكرار التراكمي (النسبي)
60 – 50	4	0.08	8	8
70 – 60	13	0.26	26	34
80- 70	21	0.42	42	76
90 – 80	7	0.14	14	90
100 – 90	5	0.10	10	100
المجموع	50	1.00	100	-

- **التكرار المتجمع الصاعد والهابط**

يهتم هذا النوع من التوزيعات بتحديد القيم التي تقل او تزيد عن قيمة معينة مقابل كل فئة من فئات التوزيع وتكون التوزيعات التكرارية المتجمعة نوعين هما:

➢ **التكرار المتجمع الصاعد**

يمكن الحصول على التكرار المتجمع الصاعد من خلال تجميع او تراكم تكرارات الجدول الاصلي بدء بتكرار الفئة الاولى وانتهاء بتكرار الفئة الاخيرة منه الى ان نحصل على مجموع التكرارات كتكرار متجمع صاعد

➢ **التكرار المتجمع الهابط**

يمكن الحصول على التكرار المتجمع النازل من خلال طرح تكرارات الجدول الاصلي من مجموع التكرارات على التوالي بدء بتكرار الفئة الاولى وانتهاء بتكرار الفئة الاخيرة منه الى ان نحصل على التكرار الاخير كتكرار متجمع نازل للفئة .

وبالعودة الى الجدول رقم (2.4) يمكن احتساب التكرار المجتمع الصاعد والهابط على النحو الاتي .

جدول (2.5)التوزيع التكراري ذو الفئات

الفئات	التكرار (f)	التكرار المتجمع الصاعد	التكرار المتجمع الهابط

60 – 50	4	4	50
70 - 60	13	17	46
80- 70	21	38	33
90 - 80	7	45	12
100 - 90	5	50	5
المجموع	50	-	-

2. عرض البيانات الاحصائية

فيما تقدم تم استعراض طرق تنظيم و تلخيص البيانات الاحصائية وعرضها جدوليا باستخدام الجداول التكرارية التي تعطي صورة شاملة و واضحة عن البيانات و توزيعاتها التكرارية . لكن عرض البيانات عن طريق الجداول التكرارية بيانيا يعطي صورة أوضح وابسط عن الظاهرة المدروسة .والرسوم البيانية] [Graphs] أكثر جاذبية للعامة و تستخدم بكثرة في الدعاية والاعلان. و الرسوم البيانية لا تضيف شيئا الى المعنى بالنسبة للبيانات ، و من وجهة نظر الاحصائيين والباحثين ليست مفيدة في التحليل الاحصائي . بالرغم من أنها تستخدم كثيرا من قبل الباحثين.

➢ **القواعد الاساسية لأنشاء الرسوم البيانية**

i. العنوان : يجب أن يحتوي الرسم البياني عنوانا مختصرا يعكس أهمية وغرض الرسم
ii. مقياس الرسم : يجب أن يرسم الشكل البياني بمقياس رسم مناسب ودقيق ثنائي او خماسي أو عشري (مثلا : 5 ، 10 ، 15 أو 10 ، 20 ، 30)
iii. الحاشية : أسفل الرسم البياني (في الحاشية) يجب أن تضاف ملاحظات ايضاحية
iv. الاحداثي الراسي والافقي : العمود الراسي والافقي يجب وصفه بوحدات القياس

2.2.1 . أنواع الرسوم البيانية

➢ **الاعمدة [Bars]**

تستخدم الاعمدة في الرسوم البيانية غالبا . وتستخدم الاعمدة لعرض البيانات للمتغيرات غير المتصلة أو النوعية . وعند استخدام الاعمدة للعرض البياني يجب مراعاة الاتي:

i. عرض الاعمدة يجب ان يكون ثابتا
ii. المسافة بين الاعمدة يجب ان تكون ثابتة وموحدة
iii. الاعمدة يمكن ان تكون افقية أو راسية

iv. عند أنشاء الاعمدة يمكن كتابة الرقم المقابل لرأس العمود (اختيارا)

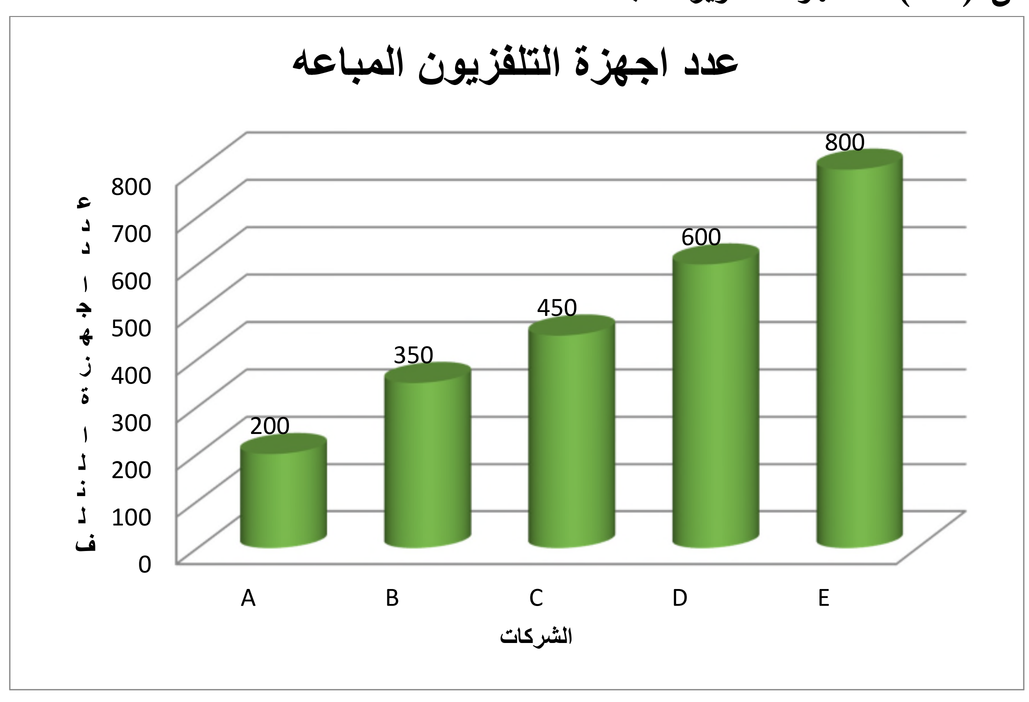
وتوجد الانواع التالية من الاعمدة

- الاعمدة البسيطة
- الاعمدة المتلاصقة
- الاعمدة المجزأة
- المدرج التكرار [Histogram]
- المضلع التكراري [Polygon]
- المنحنى التكراري [frequency curve]
- المنحنى المتجمع التكراري الصاعد والهابط

مثال(2.4) : أعرض البيانات التالية بيانيا باستخدام الاعمدة البسيطة

الشركات	A	B	c	D	E
عدد أجهزة التلفزيون المباعة	200	350	450	600	800

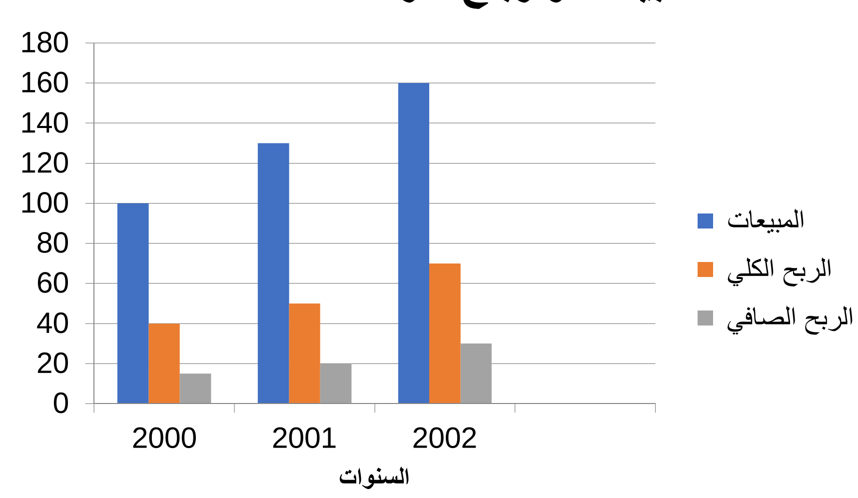
شكل (2.1) : أجهزة التلفزيون المباعة



مثال (2.5) الجدول التالي يبين المبيعات والارباح لشركة المنظفات للأعوام 2000-2002 . أستخدم الاعمدة المتلاصقة لعرض البيانات (الف ريال)

الاعوام	المبيعات	الربح الكلي	الربح الصافي
2000	100	40	15
2001	130	50	20
2002	160	70	30

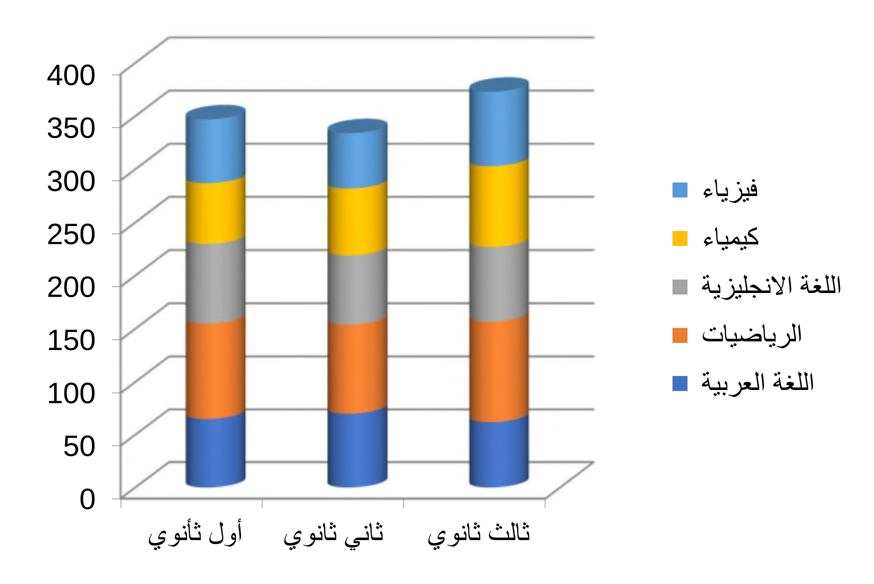
شكل 2.2 : مبيعات وارباح شركة المنظفات



مثال(2.6) أستخدم الاعمدة المجزأة للبيانات التالية التي توضح درجات 15 طالبا في الامتحانات النهائية للصف الاول ثانوي وثاني ثانوي وثالث ثانوي

الصف	اللغة العربية	الرياضيات	اللغة الانجليزية	كيمياء	فيزياء
أول ثانوي	65	90	75	57	60
ثاني ثانوي	70	84	65	63	52
ثالث ثانوي	62	95	70	76	70

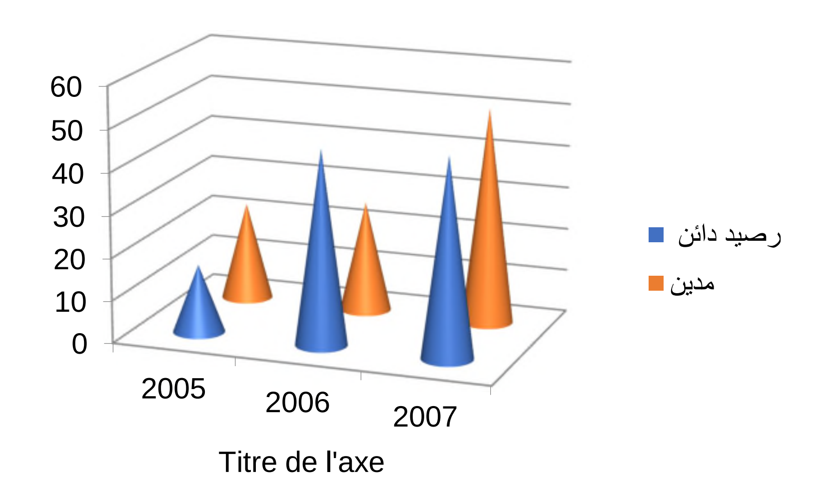
شكل 2.3 : توضح درجات 15 طالبا في الامتحانات النهائية للصف الاول ثانوي وثاني ثانوي وثالث ثانوي



مثال (2.7) الجدول التالي يبين ميزان المدفوعات لاحد البلدان (مليار دولار) للأعوام 2005-2007. أستخدم الشكل المناسب لعرض البيانات (الاعمدة)

الاعوام	رصيد دائن	مدين
2005	16	23
2006	46	26
2007	47	51

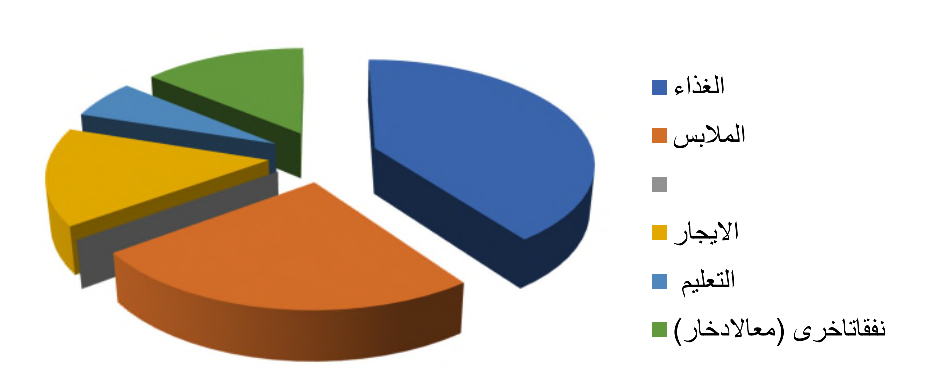
شكل (2.4) ميزان المدفوعات لبلد ما (2005-2007)



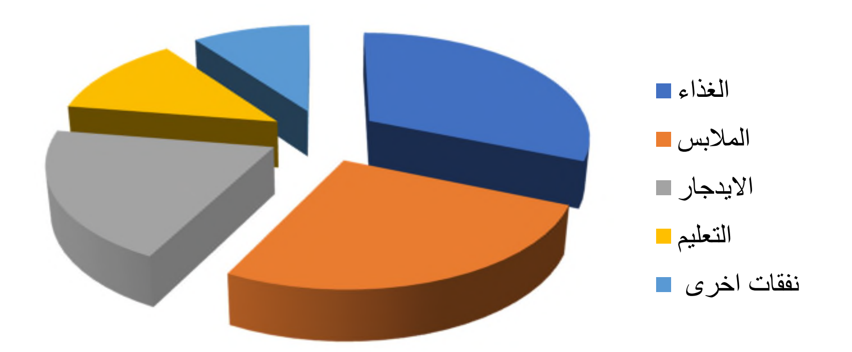
مثال (2.8) أستخدم الدوائر (pie diagram) لعرض البيانات التالية لميزانية الاسرة ($)

بنود الميزانية	الاسرة A	الاسر B
الغذاء	400	500
الملابس	250	420
الايجار	150	320
التعليم	60	200
نفقات اخرى (مع الادخار)	140	160
الاجمالي	1000	1600

شكل(2.5) ميزانية الاسرة (A)



شكل (2.6) ميزانية الاسرة (B)



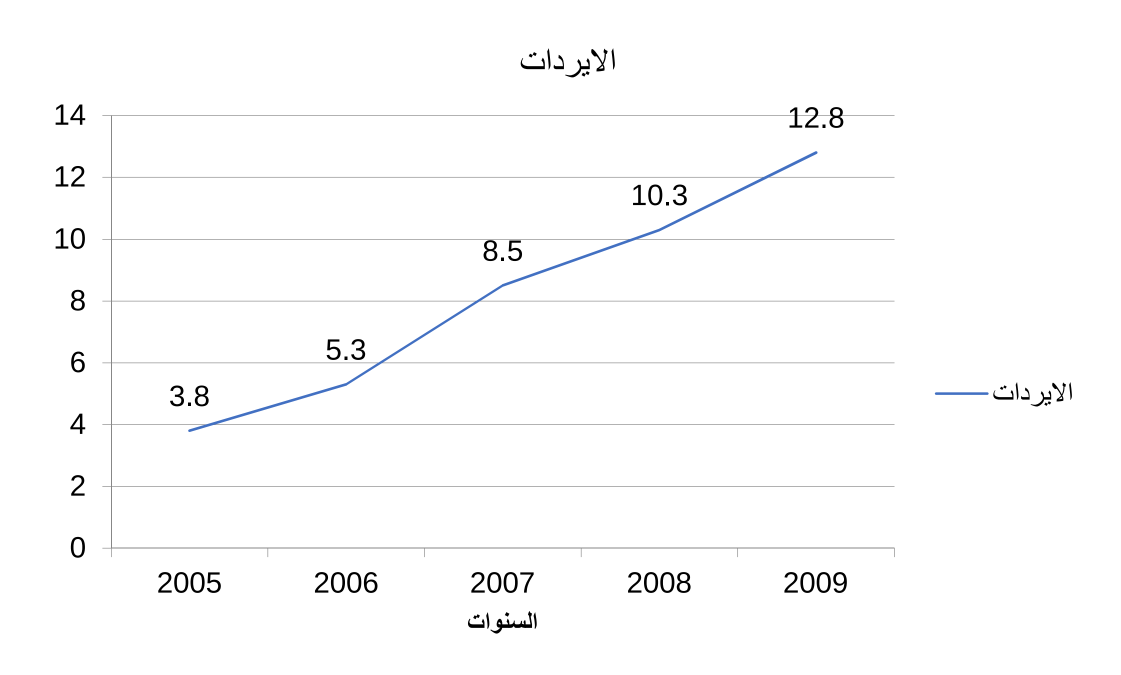
الخط البياني

مثال (2.9) أستخدم الخطوط البيانية Line Graphs لعرض البيانات الاتية

i . الجدول التالي يبين ايرادات شركة للصناعات الغذائية للأعوام 2005 - 2009 (مليون دولار)

السنوات	2005	2006	2007	2008	2009
الإيرادات	3.8	5.3	8.5	10.3	12.8

شكل (2.7) ايرادات شركة للصناعات الغذائية للأعوام 2005 - 2009 (مليون دولار)

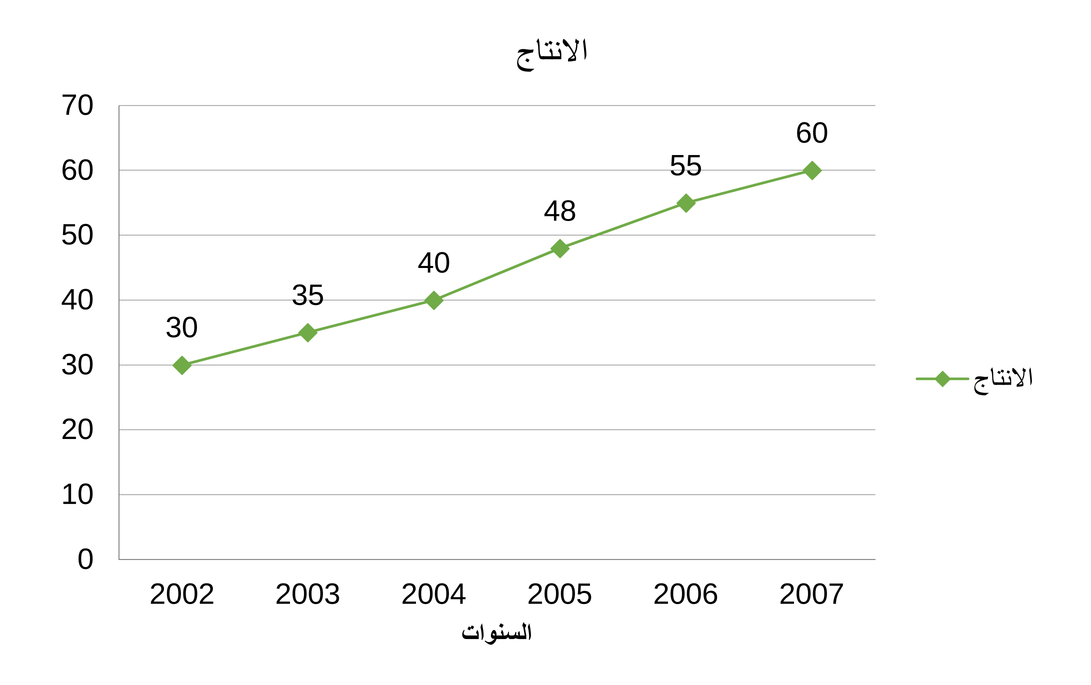


(ii) أستخدم الخط البياني لعرض البيانات الاتية لإنتاج السيارات (الف سيارة) للأعوام 2002-2008

السنوات	2002	2003	2004	2005	2006	2007
الانتاج	30	35	40	48	55	60

شكل (2.8) أنتاج السيارات (الف سيارة) للأعوام 2002-2008

الانتاج

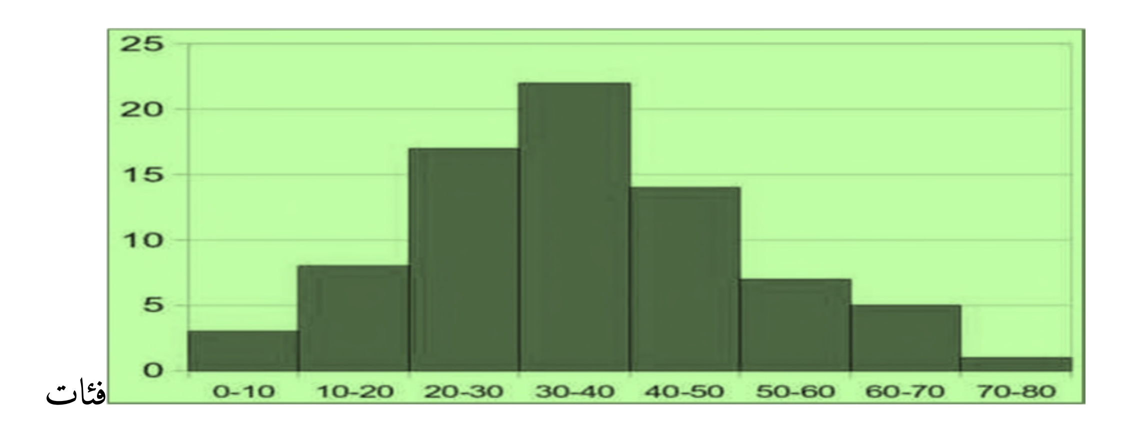


مثال (2.10) أعرض البيانات الاتية باستخدام المدرج التكراري (Histogram)

فئات العمر	10-0	-10 20	-20 30	-30 40	-40 50	-50 60	-60 70	-70 80
عدد الاشخاص	3	8	17	22	14	7	5	1

شكل (2.9) عدد الاشخاص حسب فئات العمر

الاشخاص

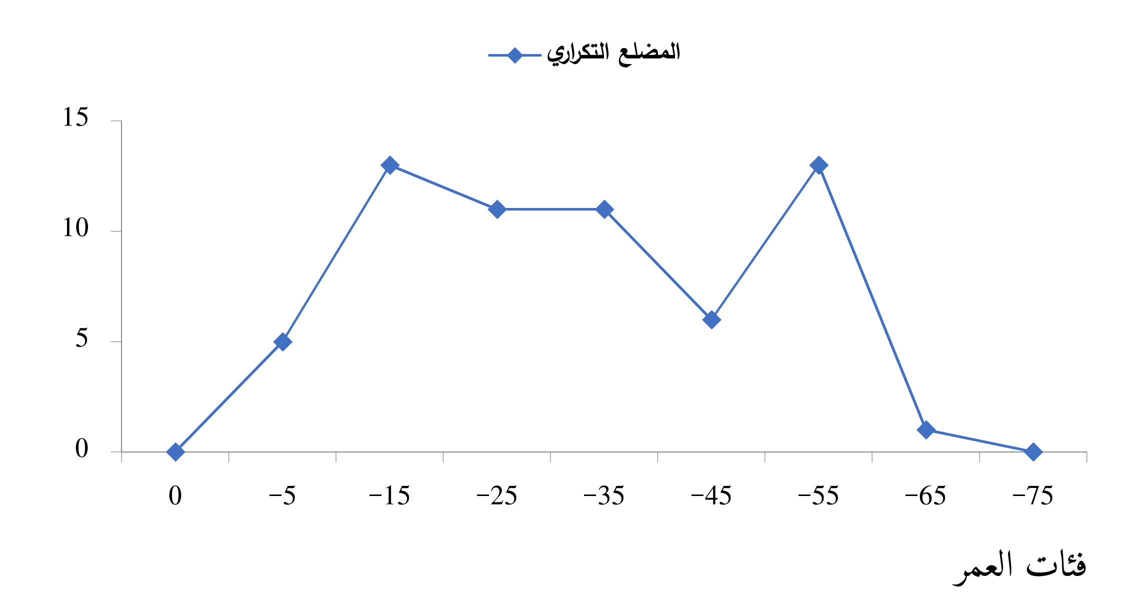


فئات العمر

مثال(2.11) أعرض البيانات الاتية باستخدام المضلع التكراري

المجموع	- 65	- 55	- 45	- 35	- 25	- 15	- 5	فئات العمر
60	1	13	6	11	11	13	5	التكرار

شكل (2.10) المضلع التكرار **التكرار**



فئات العمر

تمارين (الفصل الثاني)

1. الجدول أدناه يمثل درجات 52 طالبا في أعمال الفصل لطلاب المستوى الاول محاسبة في مساق مبادئ المحاسبة (الدرجة من 50) ، كلية العلوم الادارية.

- بوب البيانات في جدول تكراري مناسب
- أوجد التكرار النسبي والتكرار النسبي المئوي

4	3	1	2	4	5	1	2	2	3
1	1	2	3	3	2	1	2	0	3
5	5	3	4	1	2	3	0	3	2
5	6	0	1	3	4	5	6	3	4
4	3	2	1	0	2	1	3	4	5

2.الجدول أدناه يبين عدد السيارات المباعة في أحد المصانع لإنتاج السيارات خلال الفترة 2000-2005 ، أعرض البيانات باستخدام (أ) الاعمدة المنفردة (ب) الخط البياني

2004	2003	2002	2001	2000	الاعوام
1200	1700	1900	2800	2100	عدد السيارات المباعة

3.أستخدم الدائرة pie diagram) لعرض البيانات التالية

النسبة (%)	بنود الانفاق الشهري كنسبة من الدخل
20	الغذاء
12	المواصلات
15	الايجار
25	التعليم
10	الصحة
18	نفقات أخرى
100	الاجمالي

4. من الجدول التكراري أدناه أرسم

- المدرج التكراري (Histogram)
- المضلع التكراري (Polygon)
- المنحنى التكراري (Frequency curve)

الفئات	عدد الطلاب (التكرار)
10 - 0	5
20-10	16
30-20	20
40-30	28
50–40	15
60–50	10
70-60	5
80-70	2



الفصل الثالث

مقاييس النزعة المركزية





مقاييس النزعة المركزية تعني دراسة سلسلة من البيانات وفق اسس معينة للتعرف على بعض القيم المركزية التي تبين ميل مفردات السلسلة ونزعتها نحو تلك القيم المركزية . وهناك خمسة مقاييس ذات الاستخدام الشائع من قبل الاحصائيين والباحثين . وهذه المقاييس هي كالاتي:

3.1 الوسط الحسابي(Arithmetic Mean)

الوسط الحسابي هو أحد مقاييس النزعة المركزية [Central Tendency Measure] اكثرها استخداما نظرا لسهولة احتسابه وتحديده رياضيا.

ويعرف الوسط الحسابي [The Mean] لمجموعة من القيم x_1، x_2 ، $x_3......x_n$ بأنه حاصل قسمة مجموع تلك القيم على عددها . أي أن :

$$\bar{x} = \frac{x_1 ، x_2 ، x_2 ، ....، x_n}{n}$$

$$\bar{x} = \sum x / n$$

1.1.1 حساب الوسط الحسابي للبيانات غير المبوبة [un grouped data]

مثال (3.1): البيانات الاتية تبين المرتب الشهري لعشرة موظفين في البنك (الف ريال يمني) . أحسب الوسط الحسابي .

150	100	70	90	85	80	60	55	50	30

الحل :

$$\bar{x} = \frac{x_1 ، x_2 ، x_2 ، ....، x_n}{n}$$

$$\bar{x} = \sum x / n$$

$$\bar{x} = \frac{30 + 50 + 55 + 60 + 80 + 85 + 90 + 70 + 100 + 150}{10}$$

$$= 770/10 = 77$$

الوسط الحسابي لمرتبات الموظفين في البنك = 77000 ريال

3.1.2: الوسط الحسابي للبيانات المبوبة[Grouped Data]

في حالة البيانات المبوبة فأننا نستخدم القانون الاتي لإيجاد الوسط الحسابي :

$$\bar{x} = \frac{\sum f\mathrm{x}}{\sum f}$$

مثال (3،2) : أحسب الوسط الحسابي للبيانات المبوبة الاتية لدرجات طلاب المستوى الاول ، كلية العلوم الادارية ،في مساق مبادئ المحاسبة .

الفئات (الدرجات)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
التكرار	2	3	4	5	6	7	10	8	3	2

الحل : نكون جدولا يلخص خطوات الحل على النحو الاتي :

جدول(3،1) درجات طلاب المستوى الاول ، كلية العلوم الادارية ،في مساق مبادئ المحاسبة.

الفئات (درجات الطلاب)	التكرار	مركز الفئة	*Fx*
0 – 10	2	5	10
10-20	3	15	45
20-30	4	25	100
30-40	5	35	175
40-50	6	45	270
50-60	7	55	385
60-70	10	65	650

70-80	8	75	600
80-90	3	85	255
90-100	2	95	190
المجموع	50	-	2680

$$\bar{x} = \frac{\sum f\text{x}}{\sum f}$$

$\bar{x}$= 2680/50 = 53.6

3.2: الوسط الهندسي [Geometric Mean]

يعرف الوسط الهندسي لمجموعة من القيم بأنه الجذر النوني لحاصل ضرب هذه القيم ويرمز له بالرمز [G]

أذا كان لدينا القيم :

$\boldsymbol{x_1}$ ، $\boldsymbol{x_2}$ ، $\boldsymbol{x_3}$ ، ....$\boldsymbol{x_n}$

فأن الوسط الهندسي [G] يعطى بالقانون الاتي :

$$\boldsymbol{G = \sqrt[n]{x_1 ، x_2 ، x_3 ، \ldots\ldots\ldots\ldots x_n ،}}$$

ويمكن أيضا حساب الوسط الهندسي باستخدام اللوغاريتم باستخدام القانون الاتي :

Log G = 1/n [log x_1 + log x_2 + log x_3 + ….+ log x_n]

Log G = 1/n [$\sum_{i=0}^{n} xi$]

G = Anti log [1/n $\sum_{i=0}^{n} xi$]

مثال (3.3) : أوجد الوسط الهندسي للقيم الاتية " 3 ، 7 ، 8 ، 11،...

$$G = \sqrt[n]{x_1 ، x_2 ، x_3 ، \ldots\ldots\ldots\ldots x_n ،}$$

$$G = \sqrt[4]{3 \text{ x } 7\text{x } 8 \text{ x } 11}$$

$G = = \sqrt[4]{1848}$

$G = 6.5$

مثال (3،4) : أوجد الوسط الهندسي للبيانات الاتية :

6 ، 169 ، 11، 112 ، 14 ، 75 ، 35 ، 215

$Log \quad = 1/n \left[\sum_{i=0}^{n} xi \right]$

$G = Antilog \left[1/n \sum_{i=0}^{n} xi \right]$

جدول (3،2) حساب الوسط الهندسي عن طريق اللوغاريتم

X	Log_{10} x
6	0.7781
169	2.2279
11	1.0414
112	2.0492
14	1.1461
75	1.8750
35	1.3979
215	2.3324
$\sum \log_{10} x$	10.8066

$Log \ G \quad = 1/n \left[\sum_{i=0}^{n} xi \right]$

$Log\ G = 10.8066/8 = 1.3508$

$G = Antilog\ (1.3508)$

$G = 22.43$

3،3 : الوسط التوافقي [Harmonic Mean]

يعرف الوسط التوافقي لمجموعة من القيم بأنه مقلوب الوسط الحسابي لمقلوب القيم ويرمز له بالرمز (H)

$$H = \frac{(n)}{\sum_{i=}^{n} 1/xi}$$

مثال (3.5) : أوجد الوسط التوافقي للبيانات الاتية : 20 ، 70 ، 50 ، 18 ، 30 ، 45 ، 66

$$\mathbf{H} = \frac{(n)}{\sum_{i=}^{n} \mathbf{1/xi}}$$

$$H = \frac{7}{\frac{1}{20}+\frac{1}{70}+\frac{1}{50}+\frac{1}{18}+\frac{1}{30}+\frac{1}{45}+\frac{1}{66}}$$

$$H = \frac{7}{0.05+0.014+0.02+0.055+0.033+0.022+0.015}$$

$$H = \frac{7}{0.209} = 33.49$$

3.4: الوسيط [Median]

يعد الو سيط أحد مقاييس النزعة المركزية . ويعرف بأنه القيمة المحددة التي تقسم مجموعة من القيم الى نصفين متساويين . وتكون القيم التي قبله أصغر منه ، والتي تليه أكبر منه ، والعكس صحيح أذا رتبت القيم تنازليا .

- **حساب الوسيط**

أولا: في حالة البيانات الغير المبوبة

توجد حالتان لحساب الوسيط في حالة البيانات الغير مبوبة :

- أذا كان عدد القيم الغير المبوبة فرديا :

أذا كان لدينا قيم المشاهدات ، x_1 ، x_2 ، x_3 ،................x_n وكانت n فردية ،

فلحساب الوسيط نتبع الخطوات الاتية .

I. نرتب القيم تصاعديا أو تنازليا

II. نوجد ترتيب الوسيط من العلاقة الاتية :

ترتيب الوسيط = $\frac{\text{عدد المفردات} + 1}{2}$

ترتيب الوسيط = $\frac{1 + n}{2}$

III . نوجد قيمة الوسيط وهي المناظرة لترتيب الوسيط

إذا كان عدد القيم الغير المبوبة زوجيا :

ترتيب الوسيط أذا كان عدد القيم الغير المبوبة زوجيا يقع بين القيمة التي ترتيبها n/2 والقيمة التي ترتيبها n /2 + 1 . وقيمة الوسيط هي متوسط هاتين القيمتين .

ترتيب الوسيط الاول = n/2

ترتيب الوسيط الثاني = n /2 + 1 = $\frac{n+2}{2}$

مثال (3.6) اوجد الوسيط للقيم الاتية

20 ، 15 ، 3 ، 10 ، 6 ، 8 ، 12 ، 22

الحل

- **نرتب القيم تصاعديا**

3 ، 6 ، 8 ، 10 ، 12 ، 15 ، 20 ،

- **نوجد ترتيب الوسيط**

بما أن القيم فردية فأن ترتيب الوسيط :

ترتيب الوسيط = 1 + n/2

$\frac{7+1}{2}$ = 8 /2 = 4

- **نوجد قيمة الوسيط وهي القيمة التي تناظر الترتيب الرابع = 10**

مثال (3.7) : أوجد الوسيط للقيم الأتية :

32 ، 22 ، 17 ، 9 ، 13 ، 5 ، 16 ، 41 ، 50 ، 26

الحل

- **نرتب القيم تصاعديا**

5 ، 9 ، 13 ، 16 ، 17 ، 22 ، 26 ، 32 ، 41 ، 50

بما أن القيم زوجية فأن هناك وسيطين للقيم

- **نوجد ترتيب الوسيط**

ترتيب الوسيط الاول = n /2 = 10 / 2 = 5

ترتيب الوسيط الثاني = n+2/2 = (10 +2) / 2 = 6

نجد القيم المناظرة للترتيبين الخامس والسادس هما : 17 و 22

الوسيط = (17 + 22) /2 = 39/ 2 = 19.5

➢ الوسيط للبيانات المبوبة

لحساب الوسيط للبيانات المبوبة نستخدم القانون الاتي :

الوسيط = بداية الفئة الوسيطية + ترتيب الوسيط – التكرار السابق x طول الفئة الوسيطية

التكرار اللاحق – التكرار السابق

مثال (3.8) : أوجد الوسيط للبيانات التالية

الفئات	60-50	70-60	80-70	90-80	100-90	-100 110
التكرار	0	6	12	47	25	10

الحل

نكون الجدول التكراري المتجمع الصاعد ، كما هو مبين في الجدول الاتي :

جدول (3.3) فئات متجمعة صاعدة لحساب الوسيط

الفئات	التكرار	فئات متجمعة صاعدة	تكرار متجمع صاعد
60-50	0	أقل من 60	0
70-60	6	أقل من 70	6
80-70	12	أقل من 80	18 ←
90-80	47	أقل من 90	65 ←
100-90	25	أقل من 100	90

110-100	10	أقل من 110	100

نوجد ترتيب الوسيط = (مجموع التكرارات) / 2 = 100/ 2 = 50

الفئة الوسيطية = (80 -90)

بداية الفئة الوسيطية = 80

نبحث في العمود الاخير عن موقع 50 ، نجد أنها تقع بين (18 ، 65)

التكرار السابق = 18

التكرار اللاحق = 65

$$\text{ألوسيط} = \text{بداية الفئة الوسيطية} + \frac{\text{ترتيب الوسيط} - \text{التكرار السابق}}{\text{التكرار اللاحق} - \text{التكرار السابق}} \times \text{طول الفئة}$$

$$\text{الوسيط} = 80 + \frac{50 - 18}{65 - 18} \times 10$$

$$= 80 + \frac{32}{47} \times 10$$

$$= 80 + 0.68 \times 10$$

$$= 80 + 6.8$$

$$= 86.8$$

3.6 : المنوال

[Mode]

المنوال هو القيمة الأكثر تكرار (الأكثر شيوعا) في مجموعة من البيانات .

أولا : حساب المنوال من البيانات الغير المبوبة

مثال(3.9) : أوجد المنوال للقيم الأتية :

، 15 ، 13 ، 11 ،9 ، 6

الحل :

طبقا لتعريف المنوال بأنه القيمة الأكثر تكرارا ، من البيانات أعلاه ، نجد أنه أيا من القيم لم تتكرر و بالتالي لا يوجد منوال للقيم .

مثال (3.10) : أوجد المنوال

8 ، 17، 5 ، 8 ، 12 ، 7 ، 13 ، 8

الحل :

القيمة الاكثر تكرارا (شيوعا) هي 8

مثال (3.11) : أوجد المنوال

7 ، 10 ، 14 ، 16 ، 10 ، 3 ، 8 ، 14

الحل :

يوجد منوالان هما : 10 ،14 لانهما تكررا اكثر

ثانيا : حساب المنوال للقيم المبوبة

توجد طرق متعددة لإيجاد المنوال في حالة البيانات المبوبة . نستخدم هنا طريقة الرافعة ويمكن أيجاد المنوال باستخدام القانون الاتي :

$$\text{المنوال} = \text{بداية الفئة المنوالية} + \frac{(\text{التكرار اللاحق})}{(\text{التكرار السابق} + \text{التكرار اللاحق})} \times \text{طول الفئة}$$

- **خطوات أيجاد المنوال**

I. نوجد الفئة المنوالية وهي التي تقابل الفئة الاكثر تكرارا

II. نحدد بداية الفئة المنوالية

III. نحدد التكرار السابق والتكرار اللاحق لتكرار الفئة المنوالية

IV. نطبق القانون لحساب المنوال

مثال (3.12) : أوجد المنوال من البيانات الأتية

الفئات	25-20	30-25	35-30	40-35	45-40	المجموع
التكرار	3	12	21	10	4	50

جدول (3.4) حساب المنوال للبيانات المبوبة

الفئات	التكرارات
25-20	3
30-25	12
35-30 ⟵	21
40-35	10
45-40	4
المجموع	50

المنوال = بداية الفئة المنوالية + (التكرار اللاحق) / (التكرار السابق + التكرار اللاحق) x طول الفئة

المنوال = 30 + $\frac{10}{(10+12)}$ X 5

المنوال = 30 + 50 / 22 = 30 + 2.27 = 32.2

تمارين (الفصل الثالث)

1. احسب الوسط الحسابي والوسيط والمنوال من الجدول التكراري أدناه :

التكرار	الفئات
8	10 – 13
15	13 – 16
27	16 -19
51	19-22
75	22-25

54	25 -28
36	28-31
18	31-34
9	34-37
7	37-40

2.البيانات أدناه تمثل درجات 52 طالبا في أعمال الفصل في مادة اللغة الانجليزية :

I. بوب البيانات في جدول تكراري (طول الفئة = 5)

II. أوجد التكرار النسبي

III. أوجد الوسط الحسابي

21	12	5	30	32	35	15	18	19
21	42	36	30	18	19	18	8	17
28	28	26	24	25	39	30	37	35
35	15	16	10	24	21	19	17	8
	19	31	30	29	21	12	17	18
	22	22	28	25	27	29	24	26

3. أوجد الوسط الهندسي للبيانات التالية :

125 ، 1462 ، 38 ، 7، ، 0.22 ، 0.08 ، 12.75 ، 0.5

4.أوجد الوسط التوافقي للبيانات التالية ::

2574 ، 475 ، 75 ، 5، 0.8 ، 0.08 ، 0.05 ، 0.0009



الفصل الرابع

مقاييس التشتت





لاحظنا أهمية المتوسطات ، تم استعراضها في الفصل الثالث ، والتي تعطينا فكرة عامة عن مفردات البيانات المدروسة . لكن المتوسطات هي مقاييس النزعة المركزية للبيانات ولا تعطينا اي فكرة عن شكل توزيعاتها .

يقصد بالتشتت [Dispersion] لمجموعة من البيانات ، دراسة مدى التباعد أو التقارب لهذه البيانات عن وسطها الحسابي ، فكلما كانت البيانات قريبة من وسطها الحسابي كانت هذه البيانات غير مشتتة وكلما كانت البيانات بعيدة عن وسطها الحسابي كانت هذه البيانات مشتتة. ليكن لدينا المثال الاتي ، الذي يبين درجات مجموعتين من الطلاب في مساق الرياضيات :

المجموعة A	60	60	62	63	63	63	66	66	68	69
المجموعة B	10	10	50	65	70	75	75	85	100	100

نلاحظ أن مجموع الدرجات لكل من المجموعة A ، B على التوالي 640 . ومتوسط كل منهما يساوي 64 . ان البيانات للمجموعة A محصورة بين 60 و 69 . بينما في بيانات المجموعة B ، طالبين لديهما رسوب وطالبين حصلا على درجة ممتاز ، أي أن البيانات تقع بين 10 و 100 . ونلاحظ أن بيانات المجموعة الاولى اقل تشتتا مقارنة بالمجموعة B . ولدراسة هذا النوع من عدم التجانس أو التشتت أو الاختلاف . نستخدم بعض المقاييس الاحصائية ، و يمكن استعراضها كالاتي :

4.1 : المدى [Range]

هو الفرق بين أكبر قيمة و أصغر قيمة

المدى [R] = أكبر قيمة – أصغر قيمة

ويعد المدى من أبسط مقاييس التشتت واسهلها . ويعطي فكرة عامة من خلال أظهاره للمسافة الفاصلة بين طرفي السلسلة الاحصائية .

مثال (4.1) أوجد المدى لمجموعتي القيم التي تمثل درجات الطلاب في مادة الرياضيات

69 ، 60 ، 60 ، 68 ، 66 ، 63 ، 63 ، 63 ، 62 ، 66

100 ، 100 ، 0 ، 65 ، 70 ، 75 ، 75 ، 85 ، 10 ، 60

نرتب البيانات تصاعديا :

المجموعة الاولى	60	60	62	63	63	63	66	66	68	69
المجموعة الثانية	10	10	60	65	70	75	75	85	100	100

المدى = أكبر قيمة – أصغر قيمة

المجموعة الاولى (المدى) = 69- 60 = 9

المجموعة الثانية (المدى) = 100- 10 = 90

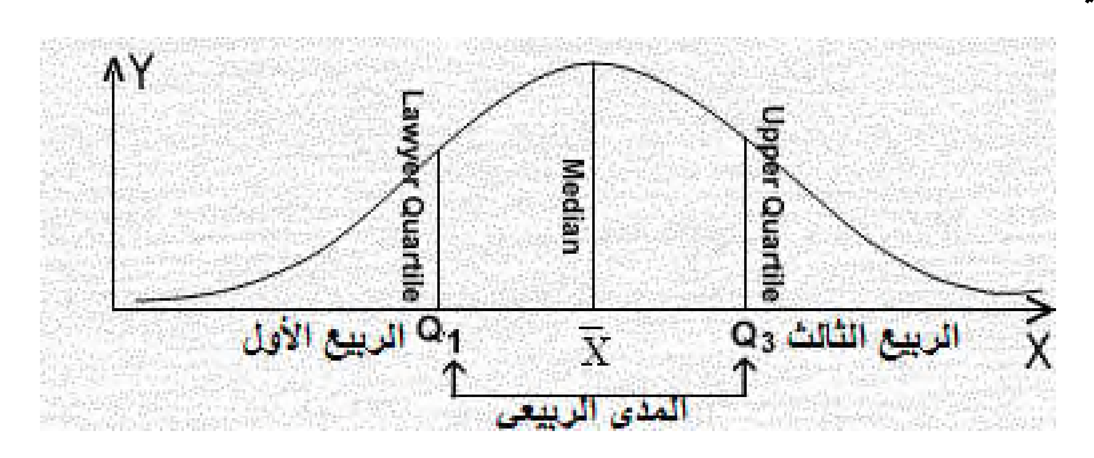
نلاحظ من المثالين أن المدى مرتبط بالقيمتين المتطرفتين . والمدى أكبر في قيم المجموعة الثانية مقارنة بقيم المجموعة الاولى . و هذا يعني أن القيم أكثر تشتتا بالنسبة للمجموعة الثانية .

4.2. نصف المدى الربيعي (Semi-Inter Quartile Range)

نصف المدى الربيعي هو أحد مقاييس التشتت المطلق و يستخدم للتغلب على العيوب الموجودة في المدى وذلك لأنه يستبعد القيم المتطرفة من الطرفين.

وهو متوسط الفرق بين الربيعيين الثالث والأول ويعرف أيضا بالانحراف الربيعي ويرمز له بالرمزQ.

ولحساب نصف المدى الربيعي ترتب البيانات تصاعدياً أو تنازلياً وتقسم الى اربعة أقسام ، نهتم منها بنهاية الربع الأول أو الأدنى (Lower Quartile) ويرمز له بالرمز Q_1 أي يستبعد الربع الأول وبداية الربع الرابع أو الأعلى (Upper Quartile) ويرمز له بالرمز Q_3 أي يستبعد الربع الرابع وهو ما نحتاج لحسابه Q_3 و Q_1 عند ربع القيم وثلاثة أرباع القيم كما يبينه الشكل الآتي:



- **خطوات حساب نصف المدى الربيعي**

i. ترتب القيم تصاعدياً أو تنازلياً

ii. إيجاد رتبة الربيع الأول (الأدنى) و رتبة الربيع الثالث (الأعلى)

باستخدام العلاقة الاتية :

رتبة الربيع الأدنى = $n+1\ /4$

رتبة الربيع الاعلى = $3(\ n+1\ )\ /4$

iii. إيجاد قيم Q_1 ، Q_3

iv. نوجد الانحراف الربيعي (نصف المدى الربيعي) باستخدام القانون الاتي :

$$\mathbf{Q = (Q_3 - Q_1)/2}$$

مثال (4.2) أوجد نصف المدى الربيعي للبيانات الواردة في المثال (1) التي تمثل درجات مجموعتين من الطلاب في مساق الرياضيات .

بيانات المجموعة الاولى : 69 ، 60 ، 60 ، 68 ، 66 ، 63 ، 63 ، 63 ، 62 ، 66

بيانات المجموعة الثانية : 100 ، 100 ، 0 ، 65 ، 70 ، 75 ، 75 ، 85 ، 10 ،60

الحل : نرتب البيانات تصاعديا

69	68	66	66	63	63	63	62	60	60
100	100	85	75	75	70	65	60	10	10

بيانات المجموعة الاولى :

رتبة الربيع الأدنى (المجموعة الاولى) :

$= (n+1)/4 = (10+1)/4 = 11/4 = 2.75$

رتبة الربيع الاعلى

$= 3\ (n+1)/4$

$= 3(10+1)/4 = 33/4= 8.25$

قيمة الربيع الادنى= قيمة الرقم الذي ترتيبه 2 + 0.75 (قيمة الفرق بين الرقم الثاني والثالث)

$Q_1 = 60 +0.75\ (\ 62-60\ ) = 60 + 1.5 = 61.5$

قيمة الربيع الاعلى = قيمة الرقم الذي ترتيبه الثامن + 0،25 (الفرق بين القيمتين الثامن والتاسع)

$Q_3 = 66 + 0.25\ (\ 86-66) = 66 + 0.5 = 66.5$

نصف المدى الربيعي (الانحراف الربيعي) Q :

$Q = (Q_3 - Q_1)/2 = (66.5 - 61.5)/2 = 5/2 = 2.5$

بالنسبة لبيانات المجموعة الثانية :

رتبة الربيع الأدنى (المجموعة الثانية) :

$= (n+1)/4 = (10+1)/4 = 11/4 = 2.75$

رتبة الربيع الاعلى

$= 3\,(n+1)/4$

$= 3(10+1)/4 = 33/4 = 8.25$

قيمة الربيع الادنى = قيمة الرقم الذي ترتيبه 2 + 0.75(قيمة الفرق بين الرقم الثاني والثالث)

قيمة الربيع الاعلى = قيمة الرقم الذي ترتيبه الثامن + 0.25 (الفرق بين القم الثامن والتاسع)

$Q1 = 10 + 0.75(60-10) = 10 + 0.75\,(50) = 10 + 37.5 = 47.5$

$Q3 = 85 + 0.25(100-85) = 85 + 0.25(15) = 85 + 3.75 = 88.75$

أذن نصف المدى الربيعي (الانحراف الربيعي) = (Q)

$Q = (Q3 - Q1)/2 = (88.75 - 47.5)/2 = 20.6$

- **نصف المدى الربيعي للبيانات المبوبة**

مثال (4.3) أوجد نصف المدى الربيعي (الانحراف الربيعي) للبيانات المبوبة الاتية :

فئات الاعمار	0 – 20	40-20	60-40	80-60	100-80
عدد الاشخاص	4	10	15	20	11

الحل

التكرار المتجمع الصاعد

فئات الاعمار	عدد الاشخاص	تكرار متجمع صاعد
0- 20	4	4
40-20	10	14
60-40	15	29
80-60	20	49
100-80	11	60

ترتيب الربيع الادنى = n / 4 = 60 / 4 = 15

ترتيب الربيع الاعلى = 3 (n / 4) = 3 (60 / 4) = 45

الفئة التي تقابل التكرار 15 = (40 – 60)

قيمة الربيع الادنى (Q_1)

$$Q_1 = L + \frac{N/4 - c.f}{f} \times i$$

$$Q_3 = L + \frac{3N/4 - c.f}{f} \times i$$

$$Q_1 = 40 + \frac{(15 - 14)}{15} X\ 20$$

$$Q_1 = 40 + 1/15\ (20)$$

$$Q_1 = 41.33$$

$$Q_3 = 60 + (45\text{-}29)/20$$

$$= 60 + 16 = 76$$

$$Q = Q_3 - Q_1 / 2$$

$$Q = (76 - 41.33)/2 = 17.33$$

$$\text{Coefficient of } Q = (Q_3 - Q_1)/(Q_3 + Q_1)$$

$$= (76\text{-}\ 41.33)/\ 76 +\ 41.33$$

$$= 34.67/117.33 = 0.29$$

معامل الانحراف الربيعي = 0.29

4.3.الانحراف المتوسط

ألانحراف المتوسط أو متوسط الانحرافات هو ناتج مجموع القيمة المطلقة (الناتج الموجب) لانحرافات القيم عن وسطها الحسابي مقسوما على عددها، ويرمز له بالرمز D_m

ويحسب الانحراف المتوسط للبيانات الغير المبوبة باستخدام القانون الاتي :

$$\mathbf{D_m = \sum | X_i - \bar{X} | / n} \qquad ، i = 1 ،2 ،3 ،4 ، .....، n$$

حيث أن :

X = قيم السلسلة

$\bar{X}$ = الوسط الحسابي لهذه القيم

مثال (4.4) أحسب الانحراف المتوسط للقيم الاتية

13 ، 15 ، 23 ، 25 ، 34

الحل

i. نوجد أولا الوسط الحسابي للقيم

$$\bar{X} = \sum X / n$$

$$= 110 / 5 = 22$$

ii. نوجد انحرافات القيم عن الوسط الحسابي

X	$X - \bar{X}$	$\lvert X - \bar{X} \rvert$
13	-9	9
15	-7	7
23	1	1
25	3	3
34	12	12
$\sum$	0	32

iii. نوجد الانحراف المتوسط

$$\mathbf{D_m} = \sum \lvert \mathbf{X_i} - \bar{X} \rvert / \mathbf{n}$$

$$D_m = 32/5 = 6.4$$

مثال (4.5) أوجد الانحراف المتوسط من البيانات التالية التي تمثل التوزيع التكراري لدرجات 50 طالبا في مساق مبادئ الإدارة .

الدرجات	0 – 10	10 -20	20 – 30	30 – 40	40 – 50	50 – 60
التكرارات	1	4	10	19	11	5

الحل

لحساب الانحراف المتوسط نستخدم القانون الاتي :

$$D_m = \sum f \; |X_i - \bar{X}| / \sum f$$

حيث أن :

$\sum f$ = مجموع التكرارات

جدول (4.1) حساب الانحراف المتوسط للبيانات المبوبة

الفئات	f	(X) مركز الفئة	f * X	$X - \bar{X}$	$\lvert X - \bar{X} \rvert$	$f \lvert X - \bar{X} \rvert$
0 – 10	1	5	5	-30	30	30
10 – 20	4	15	60	-20	20	80
20 – 30	10	25	250	-10	10	100
30 – 40	19	35	665	0	0	0
40 – 50	11	45	495	10	10	110
50 – 60	5	55	275	20	20	100
$\sum$	50	-	1750	-	-	420

$\bar{X} = \sum f X / \sum f$

$= 1750 / 50$

$= 35$

$D_m = \sum f \; |X_i - \bar{X}| / \sum f$

$= 420 / 50 = 8.4$

1.4 . الانحراف المعيار ي [Standard Deviation]

يعتبر الانحراف المعياري من أهم مقاييس التشتت و أكثرها استخداما ، نظرا لدقته ووصفه الصحيح لتشتت قيم الظاهرة المدروسة .

تعريف الانحراف المعياري : هو الجذر التربيعي للتباين ويرمز له بالرمز [s]

- **تعريف التباين [Variance]** : هو متوسط مربعات انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز [s^2]
- **حساب الانحراف المعياري**

لحساب الانحراف المعياري للبيانات الغير المبوبة نستخدم القانون الاتي :

$$\mathbf{S} = \sqrt{\frac{\sum(X-X\ )}{n}}$$

لحساب الانحراف المعياري للبيانات المبوبة نستخدم القانون الاتي :

$$S = \sqrt{\frac{\sum f\,(X-X\ )}{\sum f}}$$

مثال(4،6) أوجد الانحراف المعياري للقيم الاتية :

8 ، 12 ، 13 ، 15 ، 22

الحل

جدول (4.2) حساب الانحراف المعياري للبيانات الغير المبوبة

X	$X - \bar{X}$	$(X - \bar{X})^2$
8	-6	36
12	-2	4
13	-1	1
15	1	1
22	8	64
$\sum X = 70$	$\sum (X - \bar{X}) = 0$	$\sum(X - \bar{X})^2 = 106$

$\bar{X} = \sum X / n$

$= 70/5 = 14$

$S^2 = \sum (X - \bar{X})^2 / n$

= 106/5 = 21.2 --------------------------> [Variance]

$S = \sqrt{\sum (X - \bar{X})^2 / n}$

$= \sqrt{106/5} = \sqrt{21.2} = 4.6$ ------> Standard deviation (S.D)

مثال (4.7) من المثال رقم (4.5) أوجد التباين والانحراف المعياري و معامل الاختلاف

الحل

جدول (4.3) حساب التباين والانحراف المعياري للبيانات المبوبة

الفئات	f	مركز الفئة (X)	f * X	$X - \bar{X}$	$(X- \bar{X})^2$	$f (X- \bar{X})^2$
0 – 10	1	5	5	-30	900	900
10 – 20	4	15	60	-20	400	1600
20 – 30	10	25	250	-10	100	1000
30 – 40	19	35	665	0	0	0
40 – 50	11	45	495	10	100	1100
50 – 60	5	55	275	20	400	2000
∑	50	-	1750	-	-	6600

$\bar{X} = \frac{\sum fX}{\sum f}$

$\bar{X} = 1750 / 50 = 35$

$S^2 = \sqrt{\frac{\sum f(X-X\ )}{\sum f}}$

التباين

$S^2 = 6600/50 = 132$ -------------------> [variance]

الانحراف المعياري

$S = \sqrt{6600/50} = \sqrt{132} = 11.48$ (S.D)

معامل الاختلاف يعطى بالقانون الاتي :

$C.V = \frac{\sigma}{\bar{X}} \times 100$

حيث أن :

C.V (coefficient of variation) = معامل الاختلاف

$$C.V = \frac{\sigma}{\bar{X}} \times 100$$

$$C.V = \frac{11.48}{35} . 100 = 32.8$$

تمارين الفصل الرابع

(1) أوجد نصف المدى الربيعي (الانحراف الربيعي) ومعامل الانحراف الربيعي للبيانات المبوبة الاتية :

فئات الاعمار	-0 10	-10 20	-20 30	-30 40	-40 50	-50 60	-60 70	-70 80
التكرار	10	40	20	0	10	40	16	14

2) من البيانات ادناه

I. بوب البيانات في جدول تكراري (طول الفئة =10)
II. أوجد الانحراف المتوسط
III. أحسب التباين
IV. أحسب الانحراف المعياري
V. أوجد معامل الاختلاف

72	74	40	60	82	115	41	61	65	83
53	110	76	84	50	67	78	79	56	65
68	69	104	80	79	79	52	73	59	81
66	49	77	90	84	76	42	64	69	70
72	50	79	52	103	96	51	86	78	94

3. باستخدام معامل الاختلاف اي البيانات التالية اكثر تناسقا

A	32	28	47	63	71	39	10	60	96	14
B	19	31	48	53	67	90	10	62	40	80



الفصل الخامس

الاحتمالات [Probabilities]





1 : مفاهيم أساسية في الاحتمالات

تتناول نظرية الاحتمالات الحوادث التي تحدث كنتيجة للصدفة أو نتائج أخذت من عينات أو نتيجة أجراء تجارب .

I. **التجربة [Experiment]** : هي العملية أو مجموعة الاجراءات التي تؤدي الى نتيجة ما .

وكل تجربة تؤدي الى نتيجة أو أكثر يمكن أن نسميها حوادث . ،اذا قمنا بأجراء تجربة فأن هناك نتيجتين ممكنتين ، الاولى هي أن النتيجة تكون مؤكدة ومعروفة ، أي يمكن التنبؤ بالنتيجة بصورة مؤكدة . والثانية هي أنه لا يمكن التنبؤ بالنتيجة بصورة مؤكدة . وتسمى الظاهرة في هذه الحالة احتمالية . ومثال على ذلك في تجربة رمي حجر النرد . أذا قمنا برمي حجر النرد مرة واحدة لا نستطيع أن نتنبأ بصورة مؤكدة بأن النتيجة ستكون ظهور الرقم 6 مثلا او الرقم 4 أو 3 أو غيرها . في دراسة الاحتمالات فأن اهتمامنا يتركز على التجارب الاحتمالية حيث لا تكون نتائج التجربة معروفة سلفا .

II. **الحدث [Event]** : عند أجراء تجربة لرمي حجر النرد مرة واحدة تظهر نتيجة واحدة من بين عدة نتائج ممكنة تسمى حدثا .

مثال (5،1) تجربة رمي حجر النرد مرة واحدة تظهر النتائج الممكنة التالية

{ 1، 2 ، 3 ، 4 ، 5 ، 6 }

مثال(5،2) تجربة رمي قطعة النقود تظهر النتائج الممكنة التالية { صورة أو كتابة }

وتسمى اي نتيجة من نتائج التجربتين أعلاه حدثا .

III. **فضاء العينة [Sample space]** يمثل فضاء (فراغ) العينة مجموعة الاحداث الاولية للتجربة .

مثال (5،3) الفضاء العيني لتجربة رمي حجر النرد مرة واحدة هي

{ E_1 ، E_2 ، E_3 ، E_4 ، E_5 ، E_6 }

E_1	E_2	E_3
E_4	E_5	E_6

IV. **الاتحاد** : يشير مصطلح الاتحاد "U" إلى العملية على المجموعات التي تستخدم في دمج مجموعتين للحصول على مجموعة جديدة تحوي عناصر

مثال (5.4) اذا رمينا حجر النرد فأن الاحداث A ، B تمثل بما يلي :

A = { 4 ، 6 ، 2 }

B = { 1 ، 2 ، 3 }

A U B = { 1، 2 ، 3 ، 4 ، 5 ، 6 }

V. **التقاطع (Intersection)** : هو مجموعة العناصر المشتركة بين مجموعتين. يُشار إلى تقاطع المجموعتين بالرمز ∩

في المثال (5.5) فان تقاطع A ، B

A = { 4، 6 ، 2 }

B = { 1، 2 ، 3 }

A ∩ B = { 2 }

VI. الاحداث المتنافية [Mutually Exclusive Evens]

تكون الاحداث A و B احداثا متنافية اذا كان التقاطع A ∩ B لا يحتوي على اية احداث .

مثال (5.6) اذا كانت :

A = { 3، 4 }

B = { 7، 8 ، 9 }

A ∩ B = { Ø }

A ، B أحداث متنافية

VII. **الاحداث المتماثلة** : الاحداث المتماثلة هي الاحداث التي لها نفس الفرصة أو الاحتمال في الظهور عند أجراء تجربة ما .

مثال (5.7) عند رمي قطعة النقود مرة واحدة ، حيث الحدثين صورة أو كتابة ، تسمى أحداثا متماثلة لان احتمال ظهور الصورة يساوي احتمال ظهور الكتابة وكل منها يساوي 1/2

مثال (5.8) عند رمي قطعة النرد مرة واحدة فان هناك ستة حالات تمثل الارقام الظاهرة على حجر النرد، فان احتمال ظهور كل منهما تساوي 1/6 ، وتسمى أحداثا متماثلة

1.2 : قواعد الاحتمالات

1. احتمال حدث منفرد

اذا كان الحدث A يمكن أن يحدث بطرق عددها n_A من بين نتائج متساوية الفرصة في الوقوع عددها N . فأن احتمال A :

$$P(A) = n_A / N$$

حيث أن :

احتمال وقوع الحدث A = P(A)

n = عدد الطرق التي يمكن ان يقع بها A

N = العدد الكلي للنتائج المتساوية الفرصة في الظهور

مثال (5.9) عند القاء قطعة النقود مرة واحدة فأن الصورة والكتابة يمثلان حدثان ناتجان لهما نفس قرصة الوقوع . أي أن :

$$P(H) = n_H / N = 1/2$$

$$P(T) = n_T / N = 1/2$$

$$P(H) + P(T) = 1$$

- **نتيجة**

أذا كانت P(A) = 0 ، فأن الحدث A لا يمكن أن يقع . اذا كانت P(A) = 1 فأن الحدث A مؤكد الوقوع . و أذا رمزنا لاحتمال عدم وقوع الحدث A بالرمز P(A′) فأن:

$$\mathbf{P(A) + P(A') = 1}$$

وتتراوح قيمة P(A) بين 0 و 1

أي أن :

$$\mathbf{1 \le P(A) \le 1}$$

مثال (5.10) عند القاء حجر نرد غير متحيز مرة واحدة فأن هناك ستة نتائج متساوية الفرصة في الوقوع وهي ، 1 ، 2 ، 3 ، 4 ، 5 ، 6 و من ثم فأن :

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1$$

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$

ويكون احتمال عدم ظهور الرقم 1

$$\mathbf{P(1') = 1 - P(1)}$$

$$= 1- 1/6 = 5/6$$

$$P(1) + P(1') = 1/6 + 5/6 = 6/6 = 1$$

2. قاعدة جمع الاحتمالات

تستخدم قاعدة جمع الاحتمالات لتحديد احتمال وقوع وأحد من حدثين أو أكثر . ويتوقف حساب الاحتمال على طبيعة العلاقة بين الحوادث . فقد تكون الحوادث متنافية حيث لا توجد أمكانية لوقوع الحوادث معا (اي لا توجد تقاطع) أو غير متنافية عندما تكون هناك أمكانية لوقوعهما معا (أي يوجد تقاطع)

- **قاعدة الجمع للأحداث المتنافية**

P (A أو B) = $P (A \cup B) = P(A) + P(B)$

$$\mathbf{P (A \cup B) = P(A) + P(B)}$$

-
- **قاعدة الجمع للأحداث الغير المتنافية**

P (A أو B) = P (A) + P(B) – P (A و B)

$$P (A \cup B) = P(A) + P (B) - P (A \cap B)$$

$$\mathbf{P (A \cup B) = P(A) + P (B) - P (A B)}$$

3. قاعدة ضرب الاحتمالات

تتعلق قاعدة الضرب باحتمال وقوع حدثين أو أكثر معا . ويختلف الاحتمال فيما اذا كانت الحوادث مستقلة أو غير مستقلة .

- **قاعدة الضرب للأحداث المستقلة**

يعتبر الحدثان A و B مستقلين أذا كان وقوع A غير مرتبط بأي طريقة بوقوع B ، عندئذ الاحتمال المشترك للحدثين A و B هو :

P (A و B) = $P (A \cap B) = P (A) \cdot P(B)$

$$P(A \cap B) = P(A) \,.\, P(B)$$

- **قاعدة الضرب للأحداث الغير مستقلة**

يعتبر الحدثان غير مستقلين أذا كان وقوع أحدهما مرتبط بطريقة ما بوقوع الاخر . عندئذ

$$P(A \text{ و } B) = P(A \cap B) = P(A) \,.\, P(B/A)$$

$$P(A \cap B) = P(A) \,.\, P(B/A)$$

$$P(A \cap B) = P(A) \,.\, P(B/A)$$

والمعادلة تقرأ كالاتي : احتمال وقوع كلا من الحدثين A و B يساوي احتمال وقوع الحدث A مضروبا في احتمال وقوع الحدث B اذا علم أن الحدث A قد وقع فعلا .

مثال (5.11) اذا كان وعاء يحتوي على 10 كرات منها 3 بيضاء ، 4 صفراء ، 2 سوداء ، و واحدة حمراء . فما هي الاحتمالات التالية :

I. سحب كرة بيضاء
II. سحب كرة حمراء
III. سحب كرة سوداء أو صفراء

الحل

I. سحب كرة بيضاء

$$P(W) = n_w / N = 3/10$$

II. سحب كرة حمراء

$$P_R = n_R / N = 1/10$$

III. سحب كرة سوداء أو صفراء

بتطبيق القاعدة 2 لا نها أحداث متنافية

$$P(A \text{ أو } B) = P(A \cup B) = P(A) + P(B)$$

$$P(A \cup B) = 2/10 + 4/10 = 6/10$$

مثال (5.12) الجدول أدناه لعينة من الطلاب في المستوى الاول في الجامعة ، أستخدم الجدول في تحديد الاحتمالات التالية :

I. احتمال ان يكون الطالب عمره 18 أو 19

II. احتمال أن الطالب عمره 19 سنة فأكثر

III. احتمال ان يقل عمر الطالب عن 19 سنة

الاحتمال	التكرار	العمر	الحدث
0.05	5	18	A
0.45	45	19	B
0.35	35	20	C
0.13	13	21	D
0.02	2	22	E
1.00	100	-	المجموع

I. احتمال ان يكون الطالب عمره 18 أو 19

P (a أو b) = P (a ∪ b) = P (a) + P(b) = 0.05 + 0.45 = 0.50

II. احتمال أن الطالب عمره 19 سنة فأكثر

P (b أو c أو d أو e) = P (b ∪ c ∪ d ∪ e)

= P (b) + P (c) + P(d) + P (e)

= 0.45 + 0.35 + 0.13 + 0.02

= 0.95

III. احتمال ان يقل عمر الطالب عن 19 سنة = 0.05

وهو متمم الاحتمال :

1 – 0.95 = 0.05

مثال(5.13) يطلق راميان على هدف . فإذا علمنا أن احتمال اصابة الرامي الاول للهدف 7/ 4 ، وأن احتمال إصابة الرامي الثاني للهدف هو 2/3 . المطلوب :

I. ما هو احتمال اصابة الهدف مرتين

II. ما هو احتمال أن يصيب أحد الراميين الهدف

III. ما هو احتمال أن لا يصيب أي منهما الهدف

الحل

I. ما هو احتمال اصابة الهدف مرتين

بما ان الحدثين مستقلان فأن :

$$P(A \cap B) = P(A) . P(B)$$
$$= 4/7 \;.\; 2/3 = 8/21$$

II. ما هو احتمال أن يصيب أحد الراميين الهدف

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 4/7 + 2/3 - 8/21 = (12 + 14 - 8) / 21 = 18/21$$

Iii . احتمال أن لا يصيب أي منهما الهدف

$P(A' \cap B') = P(A \cup B)'$ .. de Morgan's Law

$$P(A' \cap B') = P(A \cup B)' = 1 - P(A \cup B)$$
$$= 1 - 18/21 = 3/21$$

4. الاحتمال الشرطي [Conditional Probability]

احتمال حدوث حدث ما وليكن A ، أذا علمنا أن الحدث B قد حدث فعلا يسمى بالاحتمال الشرطي . ويمكن صياغة الاحتمال الشرطي رياضيا كالاتي :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad P(B) \neq 0$$

Or

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad P(A) \neq 0$$

مثال(5.14) صندوق يحتوي على 3 كرات بيضاء و 5 سوداء . تم سحب كرتين عشوائيا ، واحدة تلو الاخرى بدون ارجاع . أوجد احتمال ان تكون الكرتان المسحوبتان سوداويتين .

الحل

- احتمال سحب كرة سوداء في المحاولة الاولى

$$P(A) = n_A / N$$
$$= 5/ 3+5 = 5/8$$

- احتمال سحب كرة ثانية سوداء ، علما أن الكرة الاولى المسحوبة سوداء

$P(B/A) = 4 / 3+4 = 4/7$

- احتمال أن الكرتين المسحوبتين سوداويتان

$P(AB) = P(A) . P(B/A)$

$= 5/8 . 4/7 = 20/56 = 5/14$

مثال (5.15) في امتحان القبول للطلاب في كلية العلوم الادارية ، 0.15 من الطلاب يرسبون في مساق الرياضيات و 0.25 في مساق اللغة الانجليزية و 0.10 منهم يرسبون في المقررين معا . تم اختيار طالب بشكل عشوائي . المطلوب :

- أذا كان الطالب راسبا في مساق الرياضيات ، فما هو احتمال أن يكون راسبا في اللغة الانجليزية
- أذا كان الطالب راسبا في اللغة الانجليزية ، فما هو احتمال ان يكون راسبا في الرياضيات
- ما هو احتمال أن يكون الطالب راسبا في أحد المساقين

الحل

نفرض أن :

A = الطلاب الراسبون في مساق الرياضيات

B = الطلاب الراسبون في مساق اللغة الانجليزية

$\mathbf{P(A) = 0.25}$

$\mathbf{P(B) = 0.15}$

$\mathbf{P(A \cap B) = 0.10}$

- احتمال ان يكون الطالب راسبا في مساق الرياضيات ،علما بأنه راسب في اللغة الانجليزية

$P(A/B) = P(A \cap B) / P(B)$

$= 0.10/0.15 = 2/3 = 0.66$

- احتمال ان يكون راسبا في مساق اللغة الانجليزية ،علما بأنه راسب في الرياضيات

$$P(B/A) = P(A \cap B) / P(A)$$
$$= 0.10/0.25 = 2/5 = 0.40$$

- احتمال أن يكون راسبا في أحد المساقين :

$$P(A \cup B) = P(A) + P(B) - = P(A \cap B)$$
$$= 0.25 + 0.15 - 0.10 = 0.30$$

5. نظرية بييز [Bays Theorem]

من قاعدة الضرب للأحداث غير المستقلة ، نجد أن :

$$P(B \cap A) = P(B) . P(A/B)$$

بقسمة الطرفين على P(B)

$$P(B \cap A)/P(B) = \frac{P(B).\ P(A/B)}{P(B)}$$
$$P(B \cap A)/P(B) = P(A/B)$$
$$P(A/B) = P(A \cap B)/P(B) \ \ldots\ldots\ldots\ldots\ldots \{ P(A \cap B) = P(B \cap A) \}$$
$$P(A/B) = P(A) . P(B)/P(B) \ \ldots\ldots\ldots\ldots \{ P(A \cap B) = P(A) . P(B) \}$$

$$\mathbf{P(A/B)} = \frac{\mathbf{P(A).\ P(B)}}{\mathbf{P(B)}} \ \ldots\ldots\ldots\ldots\ldots \{ \textbf{Bays Theorem} \}$$ ....نظرية بييز

مثال (5.16) أحدى المؤسسات البحثية لديها 200 باحث موزعين حسب الجدول ادناه

العمر	حاملو درجة البكالوريوس	حاملو درجة الماجستير	أجمالي
< 30	20	5	25
30-40	45	5	50
40+	10	15	25
الاجمالي	75	25	100

اذا تم اختيار باحث عشوائيا من المؤسسة البحثية ، أوجد

I. احتمال ان الباحث لديه درجة البكالوريوس فقط
II. احتمال أن الباحث يحمل درجة الماجستير وعمره اكبر من 40 عاما
III. احتمال أن عمر الباحث أقل من 30 عاما ويحمل درجة البكالوريوس

الحل

نفرض ان الباحث يحمل درجة البكالوريوس فقط = a

الباحث لديه درجة الماجستير= b

الباحث اقل من 30 سنة = c

الباحث أكبر من 40 سنة = d

- احتمال ان الباحث لديه درجة البكالوريوس فقط

$P(a) = 75/100 = 0.75$

- احتمال أن الباحث يحمل درجة الماجستير وعمره اكبر من 40 عاما

$P(b/d) = P(b\cap d) / P(d) = (15/100)/(25/100)$

$= 15/25 = 3/5$

- احتمال أن عمر الباحث أقل من 30 عاما ويحمل درجة البكالوريوس

$P(c/a) = P(c \cap a) / P(a)$

$= (20/100)/(75/100)$

$= 20/75 = 4/15$

تمارين الفصل الخامس

1. أذا كان وعاء يحتوي على 10 كرات متماثلة باستثناء أن منها 5 حمراء ، 3 زرقاء ، 2 خضراء ، ما هو احتمال الحصول على :

I. كرتان زرقاء من الوعاء ، عند السحب مرتين بدون ارجاع

II. كرتان خضراء ، عند السحب مرتين بدون ارجاع

III. كرتان زرقاء عند السحب مرتين مع الارجاع

1. اذا كان وعاء يحتوي على كرات 4 بيضاء ، 3 صفراء ، 2 سوداء ، وواحدة حمراء ، فما هي الاحتمالات التالية (بدون ارجاع)

I. سحب كرة صراء ثم سوداء

II. سحب كرة بيضاء ثم صفراء

III. سحب كرة صفراء ثم سوداء ثم حمراء

2. يحتوي وعاء على 10 كرات متماثلة ، 5 حمراء ، 3 زرقاء ، 2 خضراء ، سحبت كرة من الوعاء ، ما هو احتمال أن تكون ،

أ) حمراء

ب) زرقاء

ج) خضراء

د) ليست زرقاء

ه) ليست خضراء

و) خضراء او ليست خضراء

ز) ما هو احتمال سحب كرتين زرقاوين في سحبتين متتاليتين مع الارجاع .

3. في احدى المناطق الزراعية وجد ان 40% من المزارع تستخدم الري بالتنقيط و 10% تستخدم البيوت البلاستيكية ، وهي نشاطات غير متنافية ، فاذا اختيرت أحدى المزارع ، ما هو احتمال انها تستخدم الري بالتنقيط أو البيوت البلاستيكية أو كلاهما .

4. حجر النرد رميت مرتين . أوجد احتمال الحصول على الرقم 4، 5 أو 6 في الرمية الاولى و 1، 2، 3، أو 4 في الرمية الثانية.



الفصل السادس

الارتباط Correlation





فيما سبق من الفصول ، تم تقديم الطرق الاحصائية المتعلقة بمتغير واحد [x] وتوزيعه . و هناك العديد من المشكلات في الاحصاء تتضمن متغيرات متعددة . في بعض المشكلات الاقتصادية والاجتماعية تتم دراسة عدة متغيرات لدراسة العلاقة بينهما . و في البعض الاخر يتم الاهتمام بمتغير واحد و تتم دراسة المتغيرات الاخرى بعلاقتها بهذا المتغير . و يعرف هذا النوع من الدراسات في الاحصاء بالارتباط .

تبرز مشكلات الارتباط عندما يتم التساؤل فيما اذا كانت هناك أية علاقة بين زوج من المتغيرات محل الدراسة . فمثلا : التساؤل حول مدى العلاقة بين الطلب على سلعة معينة والسعر، أو بين التحصيل العلمي للطالب لمقرري الاحصاء والرياضيات ، أو بين نفقات الدعاية والاعلان وحجم المبيعات في شركة معينة أو بين الدخل والانفاق أو بين ممارسة الرياضة والانقاص من الوزن أو بين تعليم المرأة و عدد الاطفال التي تنجبهم .

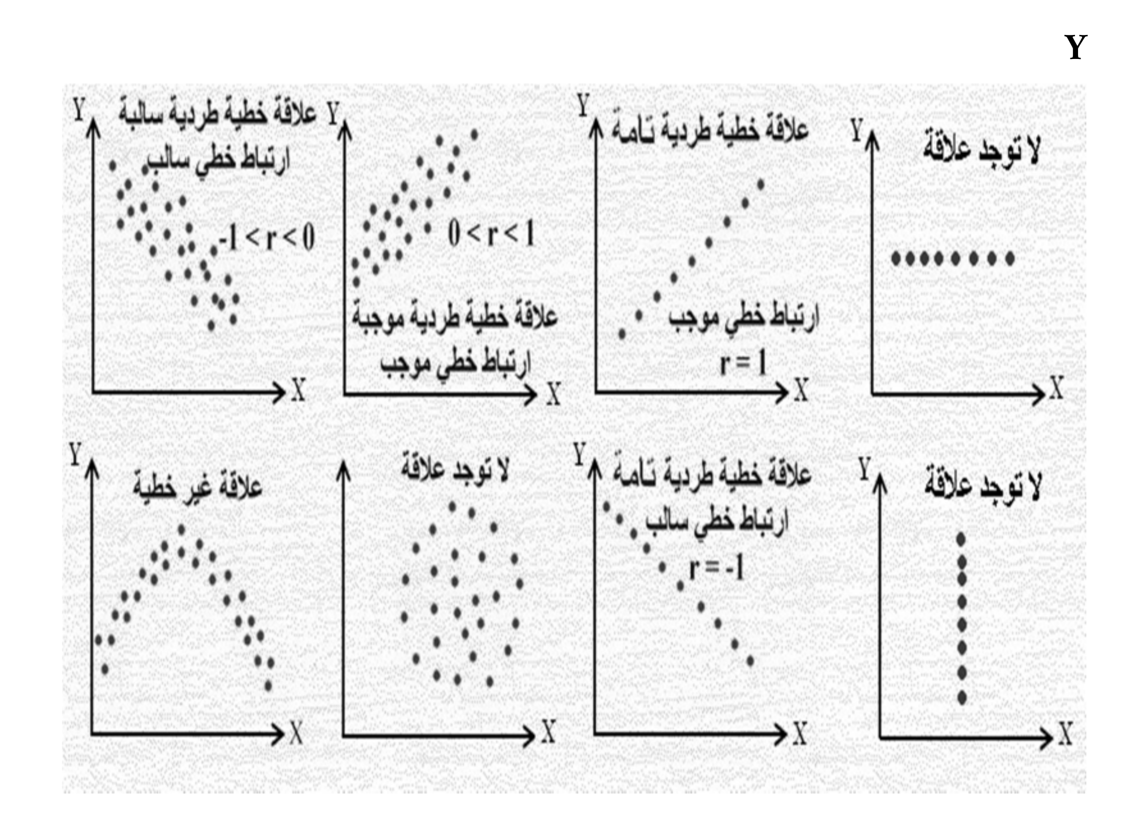
أن ظاهرة حركة متغيرين في انسجام مع بعض يسمى الارتباط . وحركة المتغيرين في انسجام وفي نفس الاتجاه تسمى ارتباط ايجابي . وعند ما يتحرك متغيرين بانسجام وفي الاتجاه المعاكس يسمى ارتباط سلبي . وعندما لا يتحرك المتغيران بانسجام مع بعض ، فأن هناك لا توجد علاقة ارتباط .

الشكل الانتشاري

إذا أخذنا (y ، x) كقيم متناظرة لمتغيرين وقمنا بتمثيلها في مستوى الإحداثيات وحصلنا على الأشكال الانتشارية ، فلكل قيمة للمتغير x توجد قيمة تقابلها للمتغير y . فالمتغير الأول يعرف بالمتغير المستقل في حين الآخر يعرف بالمتغير التابع، الشكل المرفق هنا يعرف بلوحة الانتشار وكل نقطة هنا تمثل زوج مرتب بالصورة (y،x).

شكل 6.1 : الشكل الانتشاري بين المتغيرين X و Y

Y



- **معامل الارتباط**

يعرف معامل الارتباط والذي يرمز له بالرمز [r] بأنه مقياس كمي يقيس قوة ونوع الارتباط بين المتغيرين . وتتراوح قيمته بين (+1 و -1). وتدل إشارة المعامل الموجبة على العلاقة الطردية ، بينما الاشارة السالبة على العلاقة العكسية ، وعليه فأن قيمة معامل الارتباط تتراوح بين :

$$-1 \leq r \leq +1$$

6.1 : معادلة كارل بيرسون [*Karl Pearson*] لمعامل الارتباط هي كالاتي :

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2\sum(Y-\bar{Y})^2}}$$

$X - \bar{X} = x$

$Y - \bar{Y} = y$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

حيث أن :

X ، Y = قيم المتغيرين

r = معامل ارتباط بيرسون

ويمكن حساب معامل الارتباط باستخدام الطريقة المباشرة ..وذلك باستخدام القانون الاتي :

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2}\sqrt{N\sum Y^2 - (\sum Y)^2}}$$

اذا علم الانحراف المعياري للمتغيرين وكانت n عدد المشاهدات فأن معامل الارتباط يمكن كتابته بالصيغة الأتية :

$$r = \frac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y}$$

واذا كان يمكن عدم استخدام انحرافات القيم عن الوسط الحسابي فأن معامل الارتباط يظهر كالاتي:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left\{\sum x^2 - \frac{\sum x^2}{n}\right\}\left\{\sum y^2 - \frac{\sum y^2}{n}\right\}}}$$

6.2 : معامل ارتباط الرتب [*Rank Correlation Coefficient*]

يستخدم معامل ارتباط الرتب (سيبرمان) اذا كان قياس المتغيرين ترتيبية.

أذا فرضنا ان المتغير x له الرتب R_x والمتغير y له الرتب R_y ، و ان d ترمز للفرق بين الرتبتين أي أن [$d = R_y - R_x$] فأن معامل ارتباط سيبرمان للرتب يعطى بالعلاقة الأتية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث أن n هي عدد الازواج المرتبة

مثال (6.1) أوجد معامل بيرسون للارتباط من البيانات ادناه :

عمر الزوج (X)	23	27	28	29	31	33	35	36	39
عمر الزوجة (Y)	18	22	23	24	26	28	29	30	32

الحل :

جدول 6.1 : حساب معامل ارتباط بيرسون

X	Y	$(X-\bar{X})$ x	$(Y-\bar{Y})$ y	$x\,y$	$(X-\bar{X})^2$ x^2	$(Y-Y)^2$ y^2
23	18	-8.2	-7.7	63.14	67.24	59.29
27	22	-4.2	-3.7	15.54	17.64	13.69
28	23	-3.2	-2.7	8.64	10.24	7.29
29	24	-2.2	-1.7	3.74	4.84	2.89
31	26	-0.2	0.3	-0.06	0.04	0.09
33	28	1.8	2.3	4.14	3.24	5.29
35	29	3.8	3.3	12.54	14.44	10.89
36	30	4.8	4.3	20.64	23.04	18.49
39	32	7.8	6.3	49.14	60.84	39.69
$\sum x$ =281	$\sum Y$ =232	-	-	177.46	201.56	163.91

$\bar{X} = \sum X / n = 281/9 = 31.2$

$\bar{Y} = \sum Y / n = 232/9 = 25.7$

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2\sum(Y-\bar{Y})^2}}$$

$X - \bar{X} = x$

$Y - \bar{Y} = y$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

$$r = \frac{177.58}{\sqrt{(201.56)(163.91)}} = = \frac{177.58}{\sqrt{33037.7}} = = = \frac{177.46}{\sqrt{33037.7}} = \frac{177.46}{181.7627} = 0.9763 \approx 0.98$$

اذا علم الانحراف المعياري للمتغيرين وكانت n عدد المشاهدات فأن معامل الارتباط يمكن حسابه كالاتي :

$$r = \frac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y}$$

$$S_x = \sqrt{\sum(x-\bar{x})^2/n}$$

$$= \sqrt{201.56/9} = = \sqrt{22.35} = 4.72$$

$$S_y = \sqrt{\sum(y-\bar{y})^2/n}$$

$$= \sqrt{163.91/9} = = \sqrt{18.21} = 4.26$$

$$r = \frac{\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y} = \frac{\frac{1}{9}(177.46)}{(4.72)(4.26)} = \frac{19.717}{20.10} = 0.98$$

$r = 0.98$

مثال (6.2) : أحسب معامل الارتباط من البيانات الاتية

X	13	12	10	10	8	6	6	5	5	4
Y	3	7	11	9	7	11	14	11	14	12

الحل

حساب معامل الارتباط باستخدام الطريقة المباشرة ..وذلك باستخدام القانون الاتي :

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2}\sqrt{N\sum Y^2 - (\sum Y)^2}}$$

جدول **6.2** : حساب معامل الارتباط باستخدام الطريقة المباشرة

X	Y	XY	X^2	Y^2
13	3	39	169	9
12	7	84	144	49
10	11	110	100	121
10	9	90	100	81
8	7	56	64	49
6	11	66	36	121
6	14	84	36	196
5	11	55	25	121
5	14	70	25	196
4	12	48	16	144

$\sum X=79$	$\sum Y=99$	$\sum XY=702$	$\sum X^2=715$	$\sum Y^2=1087$

$$r=\frac{N\sum XY-(\sum X)(\sum Y)}{\sqrt{N\sum X^2-(\sum X)^2}\sqrt{N\sum Y^2-(\sum Y)^2}}=$$

$$=\frac{10(702)-(79)(99)}{\sqrt{10(715)-(79)^2}\sqrt{10(1087)-(99)^2}}$$

$$r=\frac{7020-7821}{\sqrt{7150-6241}\sqrt{10870-9801}}=\frac{-801}{\sqrt{909}\sqrt{1069}}=\frac{-801}{(30.14)(32.69)}$$

$$r==\frac{-801}{985.27}=-0.813$$

مثال(6،3) : اختيرت عينة من الطلاب في كلية الاقتصاد مكونة من 10 طلاب و كانت نتائجهم في امتحان مادتي الاقتصاد والرياضيات على النحو الاتي :

أوجد معامل ارتباط الرتب .

الاقتصاد	39	65	62	90	82	75	25	98	36	78
الرياضيات	47	53	58	86	62	68	60	91	51	84

الحل

نرتب درجات الطلاب تصاعديا لكلا المادتين بإعطاء الرتبة (1) لأدنى درجة في مادة الاقتصاد وهي الدرجة (25) و أعلى رتبة للدرجة (98) وهي الرتبة 10 . وبنفس الطريقة لمادة الرياضيات حيث تعطى الرتبة (1) لادنى درجة وهي (47) و اعلى رتبة (10) لأعلى درجة وهي (91)

جدول 6،3 : حساب معامل ارتبط الرتب (سبيرمان)

الاقتصاد X	الرياضيات Y	رتبة X	رتبة Y	رتبة (x) – رتبة (Y) D	d^2
39	47	3	1	2	4
65	53	5	3	2	4
62	58	4	4	0	0
90	86	9	9	0	0
82	62	8	6	2	**4**
75	68	6	7	-1	1
25	60	1	5	-4	16
98	91	10	10	0	0
36	51	2	2	0	0
78	84	7	8	-1	1
-	-	-	-	$\sum d = 0$	$\sum d^2 = \mathbf{30}$

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6(30)}{10(100-1)}$$

$$r = 1 - \frac{180}{990} = 1 - 0.1818 = 0.818$$

الأستاذ الدكتور / علي أحمد السقاف

تمارين الفصل السادس

1.أحسب معامل بيرسون للارتباط من البيانات أدناه :

X	77	54	27	52	14	35	90	25	56	60
Y	35	58	60	40	50	40	35	56	34	42

2.البيانات التالية تبين اعمار الموظفين وعدد ايام الغياب المسجلة خلال شهر . أحسب معامل بيرسون للارتباط وفسر النتيجة

X	30	32	35	40	48	50	52	55	57	61
Y	1	0	2	5	2	4	6	5	7	8

(3) : اختيرت عينة من الطلاب في كلية الخدمة الاجتماعية مكونة من 12 طالبا وكانت نتائجهم في العمل الاجتماعي و العمل الميداني على النحو الاتي :

أوجد معامل ارتباط الرتب .

عمل اجتماعي	45	55	52	62	42	34	49	55	48	51	54	41
عمل ميداني	52	62	55	65	47	40	52	59	50	47	60	47



الفصل السابع

الانحدار [Regression]





اشرنا فيما تقدم بأن مشكلات الارتباط تبرز عندما يتم التساؤل فيما اذا كانت هناك أية علاقة بين زوج من المتغيرات أم لا . و يتم تقدير العلاقة عن طريق معامل الارتباط الذي يقيس قوة الارتباط بين متغيرين أو أكثر . لكن الارتباط لا يتمكن من المساعدة في حل مشكلات التنبؤ أو التقدير .

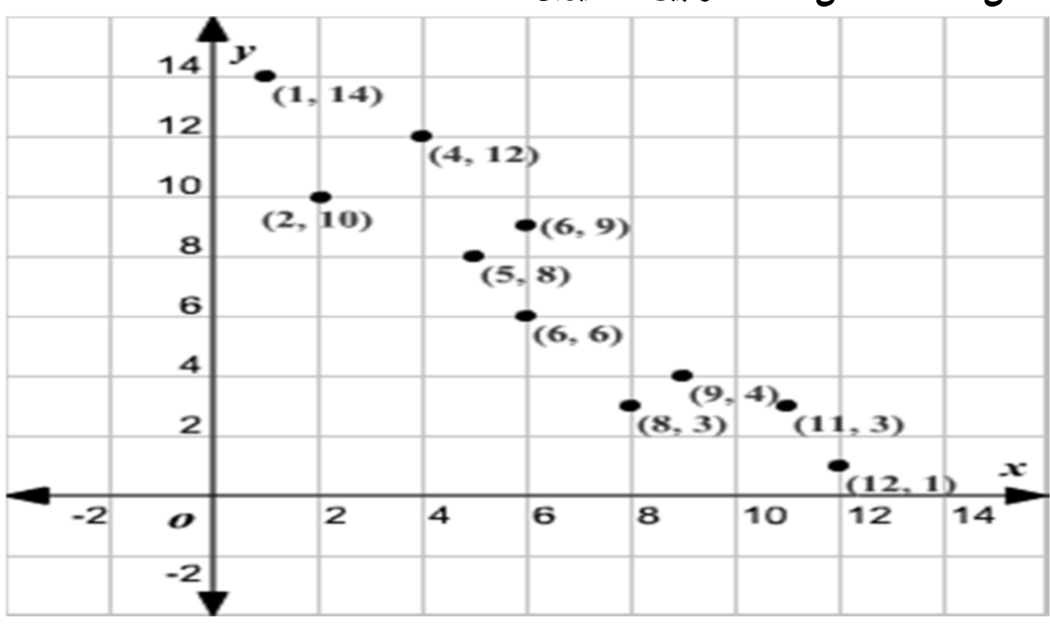
في الانحدار يتم دراسة العلاقة بين متغيرين أو اكثر على أمل استخدام العلاقة المستنتجة في المساعدة في التنبؤ أو التقدير بقيم أحد هذه المتغيرات .

و لتوضيح طريقة الانحدار الخطي البسيط والتنبؤ نفترض البيانات الأتية لقيم **X** و **y** وذلك لعينة من 10 مفردات :

X	8	2	11	6	5	4	12	9	6	1
Y	3	10	3	6	8	12	1	4	9	14

شكل الانتشار الذي يعكس العلاقة بين المتغيرين **X** و **Y** يبين في الشكل 7.1

شكل 7.1 : شكل الانتشار بين المتغيرين X ، Y



شكل (7.1) يوضح الشكل الانتشاري لهذه البيانات . و يتضح من الشكل ان العلاقة بين (**x**) و (**y**) علاقة خطية تقريبا . ولذلك فالخط المستقيم يمكن ان يوفق بين نقاط الانتشار وذلك

بغرض التنبؤ بقيم (Y) المقابلة لقيم (X) . ولأي قيمة مثلا x = 4 فأن قيمة Y المتوقعة ستكون هي المسافة الراسية الممثلة على الخط المستقيم فوق قيمة X وبقراءة قيمة y المتوقعة والمناظرة لقيمة x = 4 هي 12 تقريبا .

ولنفرض ان العلاقة بين X و Y هي علاقة خطية . ويعني ذلك انه اذا تكررت هذه التجربة عددا كبيرا من المرات (N) تحت ظروف واحدة وانه تم حساب متوسط قيم y المناظرة لكل قيمة x فأننا نحصل على مجموعة نقاط تقع تقريبا على خط مستقيم .

7.1 : طريقة المربعات الصغرى [*Least Square Method*]

كما اشرنا سلفان بان العلاقة بين x ، و Y في الشكل (7.1) يمكن ان تكون خطية أو علاقة غير خطية . و مشكلة التنبؤ الخطي تؤول الى مشكلة توفيق خط مستقيم لمجموعة من النقاط . ،احدى الطرق التي تستعمل لإيجاد ذلك الخط المستقيم هي طريقة المربعات الصغرى ، وتتلخص هذه الطريقة في أيجاد قيم معاملات المعادلة التي تجعل مجوع مربعات الاخطاء أصغر ما يمكن. ومعادلة خط المستقيم الذي يوفق البيانات في شكل الانتشار (7.1) يمكن كتابتها في الصورة:

$$Y = a + bx$$

حيث a ، b معلمتا الخط المستقيم .

وحيث أن المشكلة هي حساب قيم المعلمتين a ، b ، حتى يمكن للخط المستقيم أن يوفق مجموعة النقاط ، لذلك فأن المسالة هي أحدى المسائل لحساب قيم المعلمات بطرقة ذات كفاءة عالية . وعلى الرغم من وجود العديد من الطرق لاحتساب هذا التقدير الا أن افضل هذه الطرق لمشكلات الانحدار هي طريقة المربعات الصغرى . و حيث ان الخط المطلوب سيستخدم لأغراض التنبؤ ، لذلك فمن المناسب أن يكون ذلك الخط من الدقة بحيث تكون اخطاء التنبؤ صغيرة جدا. والمقصود هنا بأخطاء التقدير هو الفروق بين القيمة المشاهدة و القيمة المناظرة لها.

وباستخدام طريقة المربعات الصغرى نستطيع الحصول على افضل الخطوط المستقيمة التي توفق النقاط (شكل 7.1) ، والمشكلة هي ان نحصل على افضل الخطوط بطريقة منظمة ورشيدة و هذا ما يقدمه مبدأ المربعات الصغرى .

وتتلخص طريقة المربعات الصغرى في احتساب قيم تقديرية لمعالم معادلة خط الانحدار البسيط (**b ، a**) على أساس تصغير مجموع مربعات انحرافات القيم عن وسطها الحسابي .

المعادلة الخطية للانحدار البسيط هي :

$$Y = a + bx$$

فأن تقدير معلمتي المعادلة هي :

$$b = \frac{\sum xy}{\sum x^2} \qquad .....(1)$$

$$a = \overline{y} - b\,\overline{x} \qquad .....(2)$$

مثال (7.1) بالإشارة الى المثال المعطى في الشكل الانتشاري (7.1) أوجد معادلة الانحدار **Y** على **X**

الحل

الجدول 7.1 **يوضح خطوات الحسابات في طريقة المربعات الصغرى .**

i	xi	yi	$X - \overline{X}$	$Y - \overline{Y}$	$(x-\overline{X})(y-\overline{Y})$ xy	$(x-\overline{X})^2$ x^2
1	8	3	1.6	−4	−6.4	2.56
2	2	10	−4.4	3	−13.2	19.36
3	11	3	4.6	−4	−18.4	21.16
4	6	6	−0.4	−1	0.4	0.16
5	5	8	−1.4	1	−1.4	1.96
6	4	12	−2.4	5	−12	5.76
7	12	1	5.6	−6	−33.6	31.36
8	9	4	2.6	− 3	−7.8	6.76
9	6	9	−0.4	2	−0.8	0.16
10	1	14	−5.4	7	−37.8	29.16
∑	64	70	-	-	−131	118.4

$$\overline{x} = \sum x / n = 64/10 = 6.4$$

$\overline{y} = \sum y / n = 70 / 10 = 7$

$b = \frac{\sum xy}{\sum x^2}$ .....(3)

$b = \frac{-131}{118.4} = -1.1$

$a = \overline{y} - b\,\overline{x}$ .....(4)

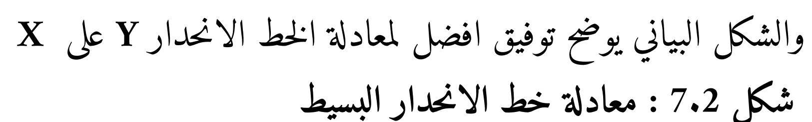
$a = 7 - (-1.1)(6.4)$

$a = 7 + 7 = 14.$

معادلة خط الانحدار :

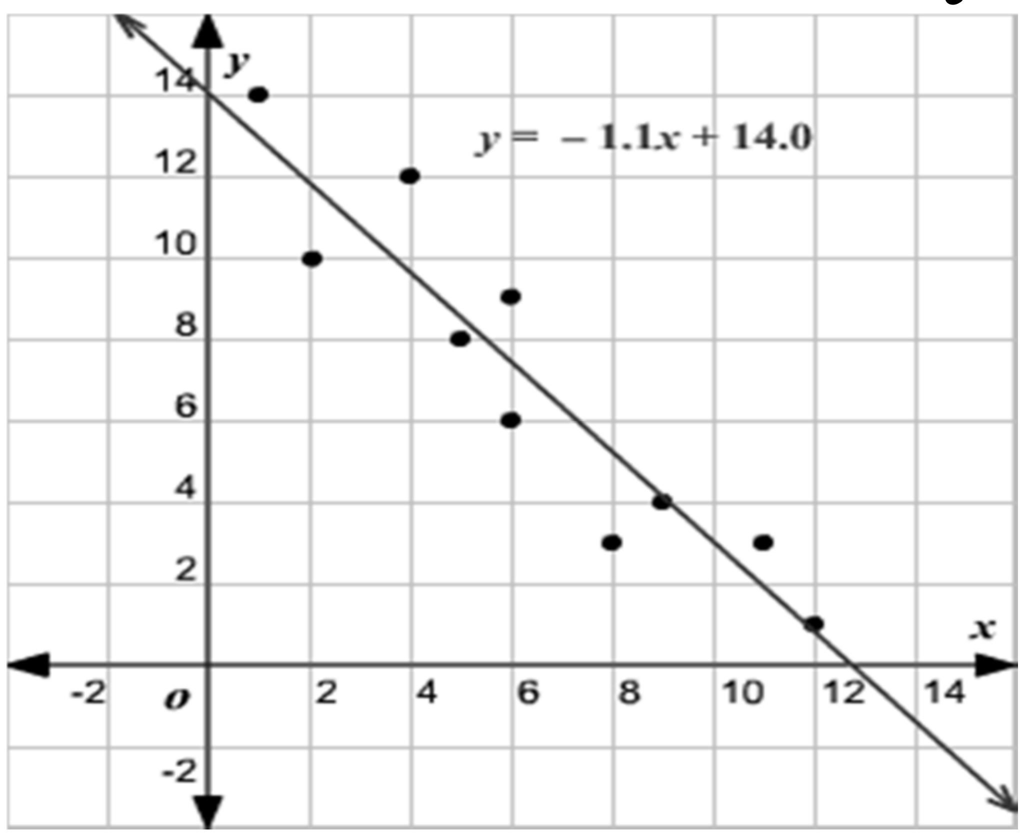
$y = a + bx$

$y = 14 - 1.1\,x$



والشكل البياني يوضح توفيق افضل لمعادلة الخط الانحدار **Y** على **X**

شكل 7.2 : معادلة خط الانحدار البسيط



في معادلة خط الانحدار البسيط التي تعطى بالعلاقة :

$Y_i = a + bx_i + e_i$

حيث أن

a و b = ثابت

X = المتغير المستقل

Y = المتغير التابع

e_i =المتغير العشوائي (حد الخطأ)

وتمثل a الجزء المقطوع من محور الصادات (قيمة y عندما تكون x = 0)

وتمثل b ميل الخط المستقيم (أي الزيادة في y اذا زادت x بمقدار وحدة واحدة .)

ويمكن حساب قيمتي b و a حسب القانونين :

$$b = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} \qquad ..... (3)$$

$$a = \bar{y} - b\bar{x} \qquad ..... (4)$$

واذا تم استخدام انحرافات القيم عن وسطها الحسابي فان تقدير قيمتي a & b من خلال المعادلتين التاليتين :

$$b = \frac{\sum xy}{\sum x^2} \qquad ......(5)$$

$$a = \bar{y} - b\bar{x} \qquad ....(6)$$

حيث أن :

$x = \bar{X} - X$

$y = \bar{Y} - Y$

مثال (7.2) من البيانات أدناه ، قدر معادلة الانحدار Y على X و X على Y

36	33	24	27	31	29	25	35	X
30	29	20	24	21	26	27	23	Y

الحل :

جدول (7.2) تقدير معادلتي الانحدار

X	Y	X^2	Y^2	XY
35	23	1225	529	805
25	27	625	729	675
29	26	841	676	754
31	21	961	441	651
27	24	729	576	648
24	20	576	400	480
33	29	1089	841	957
36	30	1296	900	1080
$\sum X = 240$	$\sum Y= 200$	$\sum X^2 = 7342$	$\sum Y^2 = 5092$	$\sum XY= 6050$

$$b = \frac{N\sum XY - \sum X\sum Y}{N\sum X^2 - (\sum X)^2}$$

$\bar{x} = \sum X /n = 240/8 = 30$

$\bar{Y}= \sum Y / n = 200/8 = 25$

$$b= \frac{8(6050)-(240 \times 200)}{8(7342)-(240)^2} = \frac{48400-48000}{58736-57600} = \frac{400}{1136} = 0.352$$

$a = \bar{y} - b\,\bar{x}$

$a = 25 - (0.352)(30)$

$a = 25 - 10.56$

$a = 14.44$

$Y = a + bX$

$Y = 14.44 + 0.352\ X$

$$b = \frac{N\sum XY - \sum X \sum Y}{N\sum Y^2 - (\sum Y)^2}$$

$\bar{x} = \sum X / n = 240/8 = 30$

$\bar{Y} = \sum Y / n = 200/8 = 25$

$$b = \frac{8(6050) - (240 \times 200)}{8(5092) - (200)^2} = \frac{48400 - 48000}{40736 - 40000} = \frac{400}{736} = 0.543$$

$a = \bar{X} - b\,\bar{y}$

$a = \bar{X} - (0.543)(25)$

$a = 30 - 13.575$

$a = 16.425$

$X = a + bY$

$X = 16.425 + 0.543\ Y$

مثال (7.3) : من البيانات أدناه حول درجات عشرة طلاب في المستوى الاول كلية العلوم الادارية في مساقي الاقتصاد والاحصاء (الدرجة من 50) أوجد :

I. معادلتي الانحدار Y على X و X على Y
II. معامل الارتباط بين درجات الاقتصاد ودرجات الاحصاء
III. الدرجة المتوقعة في الاحصاء اذا كانت الدرجة في الاقتصاد = 30

.IV

درجات الاقتصاد	25	28	35	32	31	36	29	38	34	32
درجات الاحصاء	43	46	49	41	36	32	31	30	33	39

الحل

معادلتا الانحدار :

Y = a + bx

X = a + by

$$b = \frac{\sum xy}{\sum x^2}$$

$a = \overline{y} - b\,\bar{x}$

جدول(7.3) تقدير معادلتي الانحدار لدرجات الطلاب في مساقي الاقتصاد والاحصاء

X	Y	(X-$\overline{X}$) x	(Y-$\overline{Y}$) y	x^2	y^2	xy
25	43	-7	5	49	25	-35
28	46	-4	8	16	64	-32
35	49	3	11	9	121	33
32	41	0	3	0	9	0
31	36	-1	-2	1	4	2
36	32	4	-6	16	36	-24
29	31	-3	-7	9	49	21
38	30	6	-8	36	64	-48
34	33	2	-5	4	25	-10
32	39	0	1	0	1	0
$\sum$X = 320	$\sum$Y =380	0	0	$\sum x^2 = 140$	$\sum y^2$=398	$\sum xy$= -93

$\bar{X} = \sum X / n = 320/10 = 32$

$\bar{Y} = \sum Y / n = 380/10 = 38$

$$b = \frac{\sum xy}{\sum x^2}$$

$$b = \frac{-93}{140} = -0.664$$

$a = \bar{y} - b\,\bar{x}$

$a = 38 - (-0.664)(32)$

$a = 38 + 21.248 = 59.248$

$Y = a + bX$

$Y = 59.248 - 0.664\ X$

$X = a + by$

$$b = \frac{\sum xy}{\sum y^2}$$

$$b = \frac{-93\sum}{398} = -\ 0.234$$

$a = \bar{X} - b\bar{y}$

$a = 32 - (-\ 0.234)(38)$

$a = 32 + 8.892 = 40.892$

$X = a + bY$

$X = 40.892 - 0.234\ Y$

$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$

$$r = \frac{-93}{\sqrt{(140)(398)}} = \frac{-93}{\sqrt{55720}} = \frac{-93}{236.05} = -0.394$$

$$Y = 59.248 - 0.664\,X$$

$$Y = 59.248 - 0.664\,(30)$$

$$Y = 59.248 - 19.92 = 39$$

مثال(7.4) البيانات التالية تمثل نفقات الدعاية والاعلان (X) وحجم المبيعات (Y) لشركة المرطبات والعصائر . أوجد معادلة الانحدار Y على X ثم قدر حجم المبيعات اذا كانت نفقات الدعاية والاعلان 20 . وكذلك معامل الارتباط بيرسون بين نفقات الدعاية والاعلان وحجم المبيعات .

X ($)	1100	500	300	900	1200	600	500	900
Y ($)	25000	13000	8000	20000	25000	12000	10000	15000

الحل : جدول(7.4) تقدير معادلة انحدار حجم المبيعات(Y) على نفقات الدعاية والاعلان (X)

X $ (100)	Y $ (1000)	$(X-\bar{X})$ x	$(Y-\bar{Y})$ y	x^2	y^2	xy
11	25	3.5	9	12.25	81	31.5
5	13	-2.5	-3	6.25	9	7.5
3	8	-4.5	-8	20.25	64	36
9	20	1.5	4	2.25	16	6
12	25	4.5	9	20.25	81	40.5
6	12	-1.5	-4	2.25	16	6
5	10	-2.5	-6	6.25	36	15
9	15	1.5	-1	2.25	1	-1.5
$\sum X = 60$	$\sum Y = 128$	0	0	$\sum x^2 = 72$	$\sum y^2 = 304$	$\sum xy = 141$

$\bar{X} = \sum X / n = 60/8 = 7.5$

$\bar{Y} = \sum Y / n = 128/8 = 16$

$$b = \frac{\sum xy}{\sum x^2}$$

$b = \frac{141}{72} = 1.958$

$a = \bar{y} - b\,\bar{x}$

$a = 16 - (1.958)(7.5)$

$a = 16 - 14.685 = 1.315$

$Y = a + bX$

$Y = 1.315 + 1.958\,X$

اذا كانت نفقات الدعاية والاعلان X = 20

$Y = 1.315 + 1.958\,X$

$Y = 1.315 + 1.958(20)$

$Y = 40.475 = \$\ 40475$

معامل الارتباط بيرسون

$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$

$r = \frac{141}{\sqrt{(72)(304)}} = \frac{141}{\sqrt{21888}} = \frac{141}{147.} = 0.959$

7.2 **: معامل التحديد [Coefficient of Determination]**

يعرف معامل التحديد بأنه عبارة عن نسبة التباين المفسر من التباين الكلي ، ويرمز له بالرمز r^2 . ويقيس معامل التحديد جودة التوفيق لخط الانحدار

Goodness of fit

$$\sum(Y_i - \bar{Y})^2 \quad = \quad \sum(\hat{Y} - \bar{Y})^2 \quad + \quad \sum e_i^2$$

التباين الكلي = التباين المفسر + التباين غير المفسر

[Total variation] = [Explained Variation] + [Unexplained Variation]

SST = SSR + SSE

$$\sum(Y_i - \bar{Y})^2 \quad = \quad \sum(\hat{Y}_i - \bar{Y})^2 \quad + \quad \sum e_i^2$$

$$\frac{\sum(Yi - \bar{Y})^2}{\sum(Yi - \bar{Y})^2} = \frac{\sum(\hat{Y}i - \bar{Y})^2}{\sum(Yi - \bar{Y})^2} + \frac{\sum ei^2}{\sum(Yi - \bar{Y})^2}$$

$$1 = R^2 + \frac{\sum ei^2}{\sum y^2}$$

$$R^2 = 1 - \frac{\sum ei^2}{\sum y^2}$$

مثال(7.5) من بيانات المثال رقم (7.4) حول نفقات الدعاية والاعلان (X) وحجم المبيعات (Y)أوجد معامل التحديد .

الحل : جدول(7.5) تقدير معامل التحدي لبيانات نفقات الدعاية والاعلان (X) وحجم المبيعات (Y)

X $ (100)	Y $(1000)	(X-X̄) x	(Y-Ȳ) y	x^2	y^2	xy	Ŷ	E Y- Ŷ	e^2
11	25	3.5	9	12.25	81	31.5	22.853	2.147	4.609

5	13	-2.5	-3	6.25	9	7.5	11.105	1.895	3.591
3	8	-4.5	-8	20.25	64	36	7.189	0.811	0.657
9	20	1.5	4	2.25	16	6	18.937	1.063	1.129
12	25	4.5	9	20.25	81	40.5	24.811	0.189	0.035
6	12	-1.5	-4	2.25	16	6	13.063	-1.063	1.129
5	10	-2.5	-6	6.25	36	15	11.105	-1.105	1.221
9	15	1.5	-1	2.25	1	-1.5	18.937	-3.937	15.49
$\sum X = 60$	$\sum Y = 128$	0	0	$\sum x^2 = 72$	$\sum y^2 = 304$	$\sum xy = 141$	128		27.861

$\hat{Y} = 1.315 + 1.958 X$

$\hat{Y} = 1.315 + 1.958(11) = 1.315 + 21.538 = 22.853$

$\hat{Y} = 1.315 + 1.958(5) = 1.315 + 9.79 = 11.105$

$\hat{Y} = 1.315 + 1.958\ (3) = 1.315 + 5.874 = 7.189$

$\hat{Y} = 1.315 + 1.958\ (9) = 1.315 + 17.622 = 18.937$

$\hat{Y} = 1.315 + 1.958\ (12) = 1.315 + 23.496 = 24.811$

$\hat{Y} = 1.315 + 1.958\ (6) = 1.315 + 11.748 = 13.063$

$$R^2 = 1 - \frac{\sum ei^2}{\sum y^2}$$

$$R^2 = 1 - \frac{27.861}{304} = 1 - 0.0916 = 0.908$$

$$R^2 = \frac{\sum(\hat{Y}i - \bar{Y})^2}{\sum(Yi - \bar{Y})^2} = \frac{Explained\ Vriation}{Total\ Variation} = \frac{SSR}{SST}$$

تمارين الفصل السابع

1. أوجد معادلة الانحدار من البيانات أدناه ، وقدر قيمة Y عندما تكون X = 75

99	96	90	81	78	72	69	66	63	60	X
72	70	73	78	79	82	80	84	87	85	Y

2. أوجد معادلة الانحدار Y على X ، و X على Y التي تبين درجات الطالب في مساق الاقتصاد (Y) ودرجات الطالب في مساق الاحصاء X من البيانات أدناه ، وقدر قيمة Y عندما تكون X = 75

32	34	38	29	36	31	32	35	28	25	X
39	33	30	32	32	36	41	49	46	43	Y

3. البيانات أدناه تعكس درجات 10 طلاب في الاحصاء X و الرياضيات Y

I. أحسب معامل الارتباط

II. قدر الدرجات في الرياضيات للطالب الذي حصل على 62 درجة في الاحصاء

57	69	64	60	56	57	58	58	55	56	X
66	68	66	70	68	65	70	67	67	68	Y

4. أحسب معادلة المربعات الصغرى لانحدار Y على X ثم قدر معامل التحديد R^2 من البيانات أدناه :

79	72	67	66	63	64	65	74	86	89	X
84	78	75	71	72	73	75	84	91	92	Y



الفصل الثامن

التوزيعات الاحتمالية

[Probability Distributions]





تلعب التوزيعات الاحتمالية دورا مهما في الاستدلال الاحصائي. ويتعلق بعض التوزيعات الاحتمالية بمتغيرات عشوائية منفصلة والبعض الاخر بمتغيرات عشوائية متصلة .

- **المتغير العشوائي المنفصل**: نقول عن متغير عشوائي أنه منفصل أذا كان فراغ العينة الذي يشكله المتغير يحوي عددا منتهيا قابلا للعد من النقاط .

على سبيل المثال لا الحصر، فأن المتغيرات العشوائية التي تدل على عدد البكتيريا الموجودة في واحد مليمتر مكعب من الماء وعلى عدد الاطفال الذكور لدى العائلات التي تحوي ثلاثة أطفال هي متغيرات عشوائية منفصلة وذلك لان مجموعة قيم المتغير العشوائي الاول يدل على عدد البكتيريا الموجودة في واحد مليمتر مكعب من الماء محدودة ويمكن ان تكون المجموعة 0،1 ، 2،......،[n] حيث n عددا كبيرا . وأن مجموعة قيم المتغير العشوائي الثاني الذي يدل على عدد الاطفال الذكور لدى العائلات التي تحوي ثلاثة أطفال هي المجموعة المنتهية 0، 1 ، 2،[3]

مثال(8،1) : يبين الجدول قانون التوزيع الاحتمالي للمتغير العشوائي المنفصل X الذي يدل على عدد الطلاب الذين يدخلون مختبر الحاسوب دون تحضير مادة التطبيق.

5	4	3	2	1	0	x
0.09	?	0.4	0.3	0.1	0.01	F(x)

I. أوجد $f(4)$

II. أوجد $P(x \le 3)$ ، $P(x < 3)$

III. أوجد $P(x \ge 4)$

الحل

بما أن المتغير منفصل ، فأن احتمال كل قيمة من هذه القيم أكبر أو يساوي الصفر ومجموع الاحتمالات يساوي الواحد الصحيح .

- $f(4) = ?$

$f(0) + f(1) + f(2) + \ldots\ldots + f(5) = 1$

$0.01 + 0.1 + 0.3 + 0.4 + f(4) + 0.09 = 1$

$f(4) = 1.0 - 0.9$

$f(4) = 0.1$

- $P(x \le 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$
  $= f(0) + f(1) + f(2) + f(3)$
  $= 0.01 + 0.1 + 0.3 + 0.4$
  $= 0.81$
- $P(x < 3)$
  $P(x < 3) = P(x=0) + P(x=1) + P(x=2)$
  $= 0.01 + 0.1 + 0.3$
  $= 0.41$
- $P(x \ge 4) = P(x=4) + P(x=5)$
  $= 0.1 + 0.09 = 0.19$

- **المتغير العشوائي المتصل** : نقول عن متغير عشوائي أنه متصل اذا كان يفترض أية قيمة في مجال ما أو في مجالات للأعداد الحقيقية و أن احتمال أن يأخذ المتغير قيمة معينة يساوي صفر . وهذا يعني ان مجموعة قيمه هي مجموعة غير قابلة للعد .

بما أن مجموعة المتغير العشوائي المتصل هي عبارة عن مجال او مجالات من الاعداد الحقيقية فأن مجموعة القيم هذه تشكل دالة احتمالية تصف سلوك هذا المتغير العشوائي المتصل ، أي نستطيع ان نعرف المتغير العشوائي المتصل رياضيا على النحو الاتي :

نقول عن المتغير العشوائي X بأنه متصل اذا وجدت دالة [f_x] غير سالبة معرفة ومتصلة على مجموعة الاعداد الحقيقية بحيث :

$$P(\alpha \le X \le \beta) = \int_{\alpha}^{\beta} f(x)\, dx$$

وذلك من أجل العددين الحقيقيين α، β ، حيث $\alpha \le \beta$

تسمى الدالة $f(x)$ بدالة الكثافة الاحتمالية للمتغير العشوائي المتصل و لسهولة الكتابة نرمز لها بالرمز $f(x)$ ونقول أن $f(x)$ هي دالة كثافة ، وان هذه الدالة تتصف بالخواص التالية :

I. $f(x) \ge 0$

II. $\int_{a}^{b} f(x)dx = 1$

مثال (8.2) دالة الكثافة الاحتمالية للمتغير x تعطى :

$(f(x) = 0)$ $\qquad x < 1$

$\left(f(x) = \frac{-3x^3}{8} + \frac{3x^2}{2} - \frac{9x}{8}\right) \qquad 1 \leq x \leq 3$

$(f(x) = 0\,) \qquad x > 3$

أثبت أن المساحة تحت المنحنى يساوي الواحد الصحيح

الحل :

المطلوب اثباته :

$$\int_{-\infty}^{\infty} f\,(x)dx = 1$$

$$\int_{-\infty}^{\infty} f\,(x)dx = \int_{-\infty}^{1} f\,(x)dx + \int_{1}^{3} f\,(x)dx + \int_{3}^{\infty} f\,(x)dx$$

$$= \int_{-\infty}^{1} 0\,dx + \int_{1}^{3} \left(\frac{-3x^3}{8} + \frac{3x^2}{2} - \frac{9x}{8}\right)dx + \int_{3}^{\infty} 0\,dx$$

$$= \int_{1}^{3} \left(\frac{-3x^3}{8} + \frac{3x^2}{2} - \frac{9x}{8}\right)dx$$

$$= \left| \left(\frac{-3x^4}{32} + \frac{x^3}{2} - \frac{9x^2}{16}\right) \right|_{1}^{3}$$

$$= \left(-\frac{3.81}{32} + \frac{3}{32}\right) + \left(\frac{27}{2} - \frac{1}{2}\right) + \left(-\frac{81}{16} + \frac{9}{16}\right)$$

$$= -7.5 + 13 - 4.5$$

$$= 13 - 12 = 1$$

أن دراسة المتغيرات العشوائية المتصلة والمنفصلة والتوزيعات الاحتمالية المصاحبة لها تساعدنا في الحصول على نتائج يمكن استخدامها في تقدير معالم المجتمع كذلك اختبارات الفروض المتعلقة باتخاذ القرارات .

ونستعرض فيما يلي عددا من التوزيعات الاحتمالية المهمة التي لها العديد من التطبيقات مثل توزيع ذي الحدين ، وتوزيع بواسون وذلك للمتغير العشوائي المنفصل والتوزيع الطبيعي وتوزيع ت وتوزيع ف للمتغير العشوائي المتصل .

8.1: تويع ذي الحدين :[Binomial Distribution]

التوزيع الاحتمالي الثنائي أو **ذو الحدين** هو توزيع لتجربة عشوائية لها ناتجين فقط أحدهما نجاح التجربة والآخر فشلها ويكون الشرط الأساسي أن احتمال النجاح لا يتأثر بتكرار التجربة. مثال على ذلك : رمي قطعة نقود، الإحصاءات أو الأسئلة التي تعتمد الإجابة لا أو نعم.

خصائص التوزيع الثنائي ذي الحدين :

يتميز التوزيع الثنائي بعدة خصائص هي:

I. تتكون التجربة من أكثر من محاولة. إذا تكونت التجربة من محاولة واحدة ،فإننا في تجربة توزيع برنولي.

II. استقلال المحاولات عن بعضها البعض أي ثبات احتمال النجاح p ومن ثم احتمال الفشل q.

III. هذه المحاولات جميعا متماثلة ومستقلة.

IV. حاصل جمع احتمالي النجاح والفشل يساوي الواحد الصحيح [p+q =1] لعدد n مرات من التجربة .

الصيغة العامة لتوزيع ذي الحدين تعطى كالاتي :

$$p\{x\} = c_x^n \, p^x q^{n-x}$$

$$P(x) = {}^nC_x \, P^x \, (1 - P)^{n-x}$$

مثال(3 8.) احتمال أن يكون لوالدين طفل ذو عينين زرقاوين هو ¼ فاذا كان في الاسرة 8 أطفال فما احتمال ان يكون لنصفهم على الاقل عيون زرقاء .

الحل :

احتمال النجاح {p} = ¼

احتمال الفشل {q} = 1 - p = 1 - ¼ = ¾

n = 8

ونحصل على احتمال ان يكون نصف عدد الاطفال على الاقل عيونهم زرقاء :

$$P \geq 4 = \{ p(4) + p(5) + p(6) + (7) + p(8) \}$$

$$P(x) = {}^nC_x \, P^x \, (1 - P)^{n-x}$$

$$P(4) = \frac{8!}{4!4!} \left(\frac{1}{4}\right)^4 \left(\frac{3}{4}\right)^4 = \frac{5670}{65536}$$

$$P(5) = \frac{8!}{5!3!} \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^3 = \frac{1512}{65536}$$

$$P(6) = \frac{8!}{6!2!} \left(\frac{1}{4}\right)^6 \left(\frac{3}{4}\right)^2 = \frac{252}{65536}$$

$$P(7) = \frac{8!}{7!\,1\,!} \left(\frac{1}{4}\right)^7 \left(\frac{3}{4}\right)^1 = \frac{24}{65536}$$

$$P(8) = \frac{8!}{8!\,0\,!} \left(\frac{1}{4}\right)^8 \left(\frac{3}{4}\right)^0 = \frac{1}{65536}$$

$P \geq 4 = \{ p(4) + p(5) + p(6) + (7) + p(8)$

$= \frac{5670}{65536} + \frac{252}{65536} + \frac{24}{65536} + \frac{1}{65536}$

$= \frac{5947}{65536} = 0.090$

8.2: توزيع بواسون : [Poisson Distribution]

يعتبر توزيع بواسون من التوزيعات الاحتمالية المنفصلة (المنقطعة) . و يستخدم في كثير من التطبيقات في العلوم الطبيعية والاقتصادية والاجتماعية . فاستخداماته تشمل علوم الفيزياء ودراسة الاحياء الدقيقة و بحوث العمليات والعلوم الادارية و الاقتصادية . فتوزيع بواسون كثير الاستخدام في تفتيش ومراقبة المنتوجات المصنعة وتصنيفها . وسمي بتوزيع بواسون نسبة الى عالم الرياضيات الفرنسي **سيمون بواسون** .

- **تعريف توزيع بواسون** : نقول عن المتغير العشوائي المنقطع X ، أنه يخضع لتوزيع بواسون وسطه λ أذا كانت كثافته الاحتمالية معرفة بالعلاقة :

$$\mathbf{f(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}}, \quad x = 0, 1, 2, \ldots\ldots\ldots\ldots\ldots$$

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

الشروط اللازمة لتطبيق توزيع بواسون هي :

I. أن يكون هناك ناتجان متضادان (الفشل أو النجاح)
II. يجب أن تكون الاحداث مستقلة
III. نتيجة كل تجربة أما ان الحادث ينجح أو يفشل
IV. حجم العينة [n] كبير جدا والاحتمال [p] صغير

مثال (8.4) أوجد احتمال أن يوجد 5 فيوزات معيبة على الاكثر في صندوق يحتوي على 200 فيوز . ومن التجارب معلوم ان 2% من هذه الفيوزات معيبة $\{e^{-4} = 0.0183\}$ 6

الحل

عدد الفيوزات = 200 = (n)

احتمال الحصول على فيوز معيب = 0.02

$\lambda = np$

$= 200 \times 0.02 = 4$

نريد معرفة احتمال الحصول على 5 فيوزات معيبة على الاكثر . وهي الاحتمالات للحصول على 0 ، 1 ، 2 ، 3 ، 4 او 5 فيوزات معيبة .

$$P(x) = \frac{\lambda^{x}}{x!} e^{-\lambda}$$

الاحتمال المطلوب هو = $\{P(0) + P(1) + P(2) + P(3) + P(4) + P(5)\}$

$$P = \{P(0) + P(1) + P(2) + P(3) + P(4) + P(5)\}$$

$$= \frac{4^0}{0!}e^{-4} + \frac{4^1}{1!}e^{-4} + \frac{4^2}{2!}e^{-4} + \frac{4^3}{3!}e^{-4} + \frac{4^4}{4!}e^{-4} + \frac{4^5}{5!}e^{-4}$$

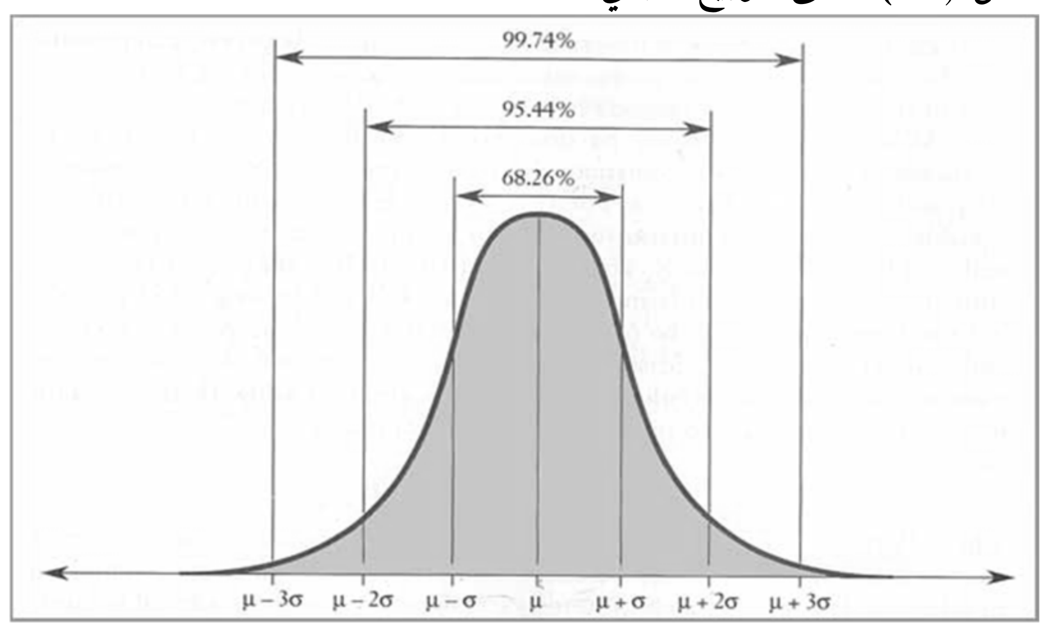
$$= e^{-4}\{1 + 4 + 8 + 32/3 + 32/3 + 128/15\}$$

$$= 0.0183\ (643/15) = 0.78$$

8.3: التوزيع الطبيعي [Normal Distribution]

يعتبر التوزيع الطبيعي من التوزيعات الاحتمالية المتصلة (المستمرة) . ويستخدم كثيرا في الواقع العملي . و شكل التوزيع يقترب من شكل الجرس . و هو متماثل (Symmetric) من حيث انتظام البيانات حول الوسط الحسابي .

شكل (8.1) منحنى التوزيع الطبيعي



- **التوزيع الطبيعي**

I. توزيع متصل له شكل الناقوس (الجرس)

II. تتساوى فيه مقاييس النزعة المركزية الوسط والوسيط والمنوال.

III. متماثل حول وسطه (صفر)

IV. الانحراف المعياري له يساوي الواحد الصحيح.

V. طرفاه يمتدان إلى مالا نهاية دون أن يلتقيا المحور الأفقي . والمساحة المهمة هي تلك الواقعة بين $\mu + \sigma$ و $\mu - \sigma$ و هي 68.26% من المساحة الكلية . وهذا يعني بأن 68.26% من المشاهدات تقع بين بين $\mu + \sigma$ و $\mu - \sigma$. والمساحة الواقعة بين $2\mu + \sigma$ و $2\mu - \sigma$ هي 95.45 % من المساحة الكلية . بينما المساحة الواقعة بين بين $3\mu + \sigma$ و $3\mu - \sigma$ هي 99.73% من المساحة الكلية

VI. المساحة تحت المنحنى الطبيعي تساوي الواحد الصحيح.

- **تعريف** : نقول عن المتغير العشوائي المستمر(المتصل) X الذي متوسطة μ وتباينه σ^2 أنه يخضع للتوزيع الطبيعي ،أذا كانت كثافته الاحتمالية معرفة كلاتي :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

التوزيع الطبيعي المعياري : [Standard Normal Distribution]

يقصد بالمنحنى الطبيعي المعياري تحويل القيم الخام الى قيم معيارية مجردة من وحدات القياس والتي يمكن الاستفادة منها في حالات المقارنة ومعرفة المناطق الواقعة تحت المنحنى الطبيعي المعياري . والصيغة الرياضية اللازمة لحساب القيمة المعيارية (Z) هي كما يلي :

$$Z = \frac{x_{i-\mu}}{\sigma}$$

مثال (8.5) أوجد المساحة تحت منحنى التوزيع الطبيعي المعياري :

Z = 1.54

Z = -1.54

- **الخطوات :**

من جدول التوزيع الطبيعي المعياري لا يجاد (Z = 1.54) نتحرك الى اسفل في عمود Z حتى نصل الى الصف 1.5 ونتحرك فيه حتى نصل الى العمود 0.04 وتقاطع الصف والعمود يعطينا القيمة المطلوبة وهي (0.4382)

Z = -1.54 = (0.4382)

مثال(8.6) أوجد المساحة تحت المنحنى الطبيعي المعياري الى يمين القيمة Z = 0.25

الحل :

نوجد اولا Z = 0.25 من جدول التوزيع الطبيعي وهي = 0.0987 . ثم نطرح القيمة الناتجة من 0.5000 (،هي نصف مساحة المنحنى) .

أي ان المساحة الى يمين Z = 0.25 هي :

0.5000 - 0.0987 = 0.4013

مثال (8.7): أوجد المساحة تحت المنحنى الطبيعي الى يسار Z = -1.60

بالبحث في جدول التوزيع الطبيعي فأن قيمة Z = - 1.60 هي (0.4452)

والمساحة الى يسار Z = - 1.60 يمكن ايجادها بطرح (0.4452) من 0.5000

0.0548 = 0.4452 – 0.5000

أي أن المساحة الى يسار z = - 0.160 هي (0.0548) أو 5.48%

مثال(8.8) أوجد المساحة تحت المنحنى الطبيعي المعياري من z = 1.60 الى z = 2.55

اولا: نوجد المساحة تحت المنحنى z = 1.60 و هي (0.4452)

ثانيا : نوجد المساحة تحت المنحنى z= 2.55 و هي (0.4946)

أذن المساحة من z = 1.60 الى z = 2.55 هي :

% 93.98 = 0.9398 = = 0.4946 + 0.4452

مثال(8.9): أذا كان عمر المصباح الكهربائي يتبع التوزيع الطبيعي بوسط حسابي = 100 وانحراف معياري = 8 . ما احتمال أن مصباحا اختير عشوائيا له عمر بين 110 و 120 ساعة احتراق .

الحل

المطلوب هو ايجاد :

$P(110 < X < 120)$

حيث أن :

$X_1 = 110$ ، $X_2 = 120$ ($\mu = 100$ ، $\sigma = 8$)

$$Z = \frac{x_{i-}\mu}{\sigma}$$

$Z_1 = 110 - 100/ 8 = 1.25$

$Z_2 = 120 - 100/ 8 = 2.50$

أي أن الاحتمال المطلوب هو المساحة بين $Z_1 = 1.25$ و $Z_2 = 2.50$

وبالبحث في جدول التوزيع الطبيعي المعياري نجد أن :

$Z_1 = 1.25 = (0.3944)$

$Z_2 = 2.50 = (0.4938)$

$Z_2 - Z_1 = 0.4938 - 0.3944 = 0.0994 = 9.94\%$

وهو الاحتمال المطلوب $P(110 < X < 120) = 0.0994$.

8.4 . توزيع [t]

ينسب توزيع t الى العالم الانجليزي جوست والذي اكتشفه عام 1908 . حيث نشر بحثا ذكر فيه معادلة هذا التوزيع ونشره باسم مستعار (Student) و اشتهر به .هذا التوزيع من حينها سمي توزيع ستيودنت [student distribution]

تعريف : التوزيع الاحتمالي الذي يطلق عليه توزيع t هو من التوزيعات الاحتمالية المستمرة (المتصلة)

ودالة كثافته هذا التوزيع تأخذ الصيغة الأتية :

$$f(t) = c\left(1 + \frac{t}{v}\right)^{-v+\frac{1}{2}} \qquad \infty < t < \infty$$

حيث أن :

v = درجات الحرية

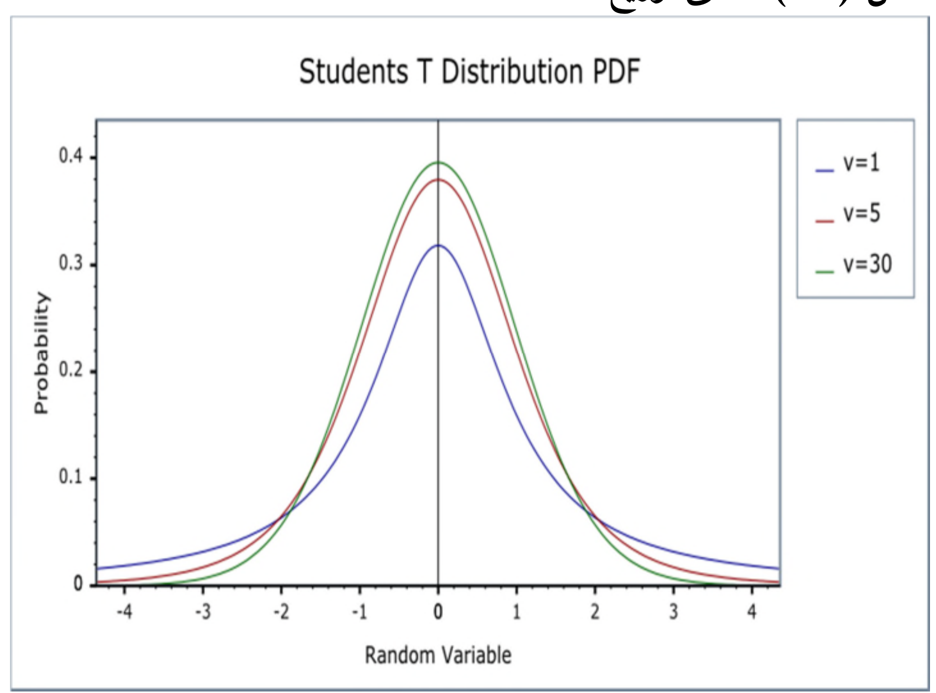
c = ثابت

والصيغة المعيارية لتوزيع t يعطى بالقانون الاتي :

$$t = \frac{\bar{x} - \mu}{\sqrt{s/n}}$$

يشبه توزيع t التوزيع الطبيعي المعياري الا انه اكثر انخفاضا منه ، ويعتمد منحنى توزيع t على معلمة هامة تحدد شكل المنحنى وهي درجات الحرية ، فعندما يزداد عدد درجات الحرية يقترب توزيع t من التوزيع الطبيعي المعياري (شكل8،2)

شكل (8،2) منحنى توزيع t



خواص توزيع t

I. المساحة الكلية تحت منحنى توزيع t تساوي الواحد الصحيح
II. يمتد طرفا المنحنى الى ما لانهاية في الاتجاهين دون ان يلامس المحور الافقي
III. منحنى توزيع t يشبه التوزيع الطبيعي المعياري فهو يشبه الناقوس (الجرس) ومتماثل حول الوسط الحسابي .

8،5 . توزيع مربع كاي (χ^2)

التوزيع الاحتمالي مربع كاي من التوزيعات الاحتمالية المستمرة (المتصلة) ودالة كثافته التوزيع تأخذ الشكل التالي:

$$F(\chi^2) = c(\chi^2)^{(v-2/2} e^{x^2}/2$$

- **خواص مربع كاي**

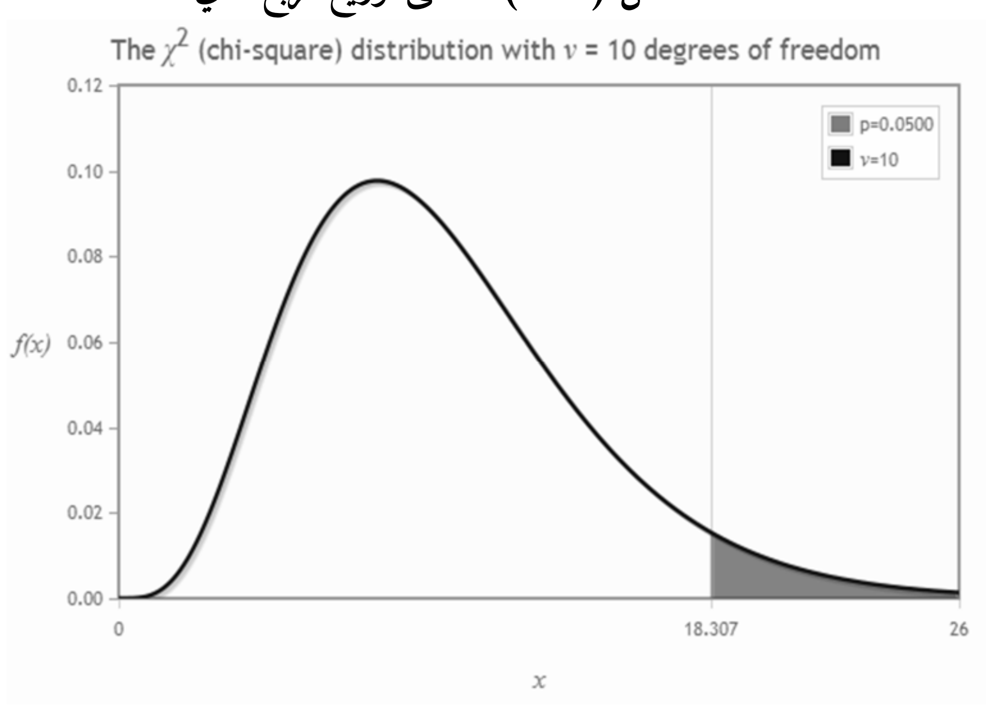
I. توزيع غير متماثل.

II. غير معرف في الجزء السالب من المستوى.

III. يبدأ من الصفر ويستمر إلى ما لانهاية.

IV. ملتوٍ ناحية اليمين أي موجب الالتواء.

شكل (8،3) منحنى توزيع مربع كاي



8،6 . توزيع F

توزيع F (فيشر) من التوزيعات الاحتمالية الهامة ويستخدم في اختبار الفرضيات.

يعطى توزيع الكثافة الاحتمالية للمتغير العشوائي F بالقانون الاتي:

$$F>0،x(F)=\frac{cF^{(V1-2)/2}}{(v2+v1F)^{v1+v2/2}}$$

خواص توزيع F

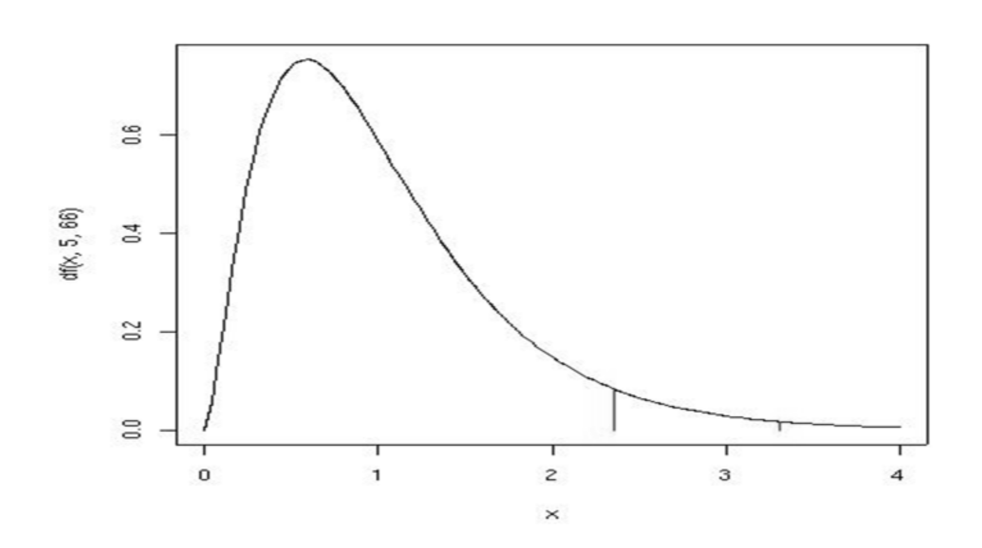
I. المساحة الكلية تحت المنحنى تساوي الواحد الصحيح

II. يعتمد توزيع **F** على معلمتين وهما درجة حرية البسط v_1 ودرجة حرية المقام v_2

III. قيم **F** موجبة دائما لا يمكن ان تكون سالبة **[F≥0]**. ويبدأ المنحنى من الصفر على المحور الافقي يمين الالتواء .

IV. منحنى توزيع **f** غير متماثل شبيه بمنحنى χ^2

شكل **(8.4**) منحنى توزيع **F**



تمارين الفصل الثامن

1. أوجد المساحة تحت منحنى التوزيع الطبيعي المعياري :

a) Z = بين $\pm$ 1.64

b) Z =بين $\pm$ 1.96

c) Z = بين $\pm$2.58

d) Z بين 0.90 و 2.10

e) Z = الى يسار 0.90

f) Z = الى يمين 2.10

2. متغير عشوائي يتبع التوزيع الطبيعي مع $\mu = 67$ و$\sigma = 3$ ما هو أحتمال ان هذا المتغير يأخذ القيمة :

أ) بين 67 و 70

ب) بين 60 و 65

ج) أقل من 60
د) أكبر من 65

3.الوسط الحسابي لإوزان مجموعة كبيرة من الناس هو 180 رطلا والانحراف المعياري . أذا كان الاوزان تتبع التوزيع الطبيعي ، أوجد احتمال أن شخصا تم اختياره عشوائيا من المجموعة سوف يزن :

أ) بين 160 و 180 رطلا
ب) اعلى من 200 رطل
ج) اقل من 150 رطل



الفصل التاسع

التقدير [Estimation]





أشرنا في الفصل الاول بأن علم الاحصاء ينقسم الى قسمين هما: الاحصاء الوصفي والاحصاء الاستدلالي . وذكرنا بأن الاحصاء الوصفي يحتوي الاساليب (الطرق) المستخدمة لتلخيص ووصف البيانات الرقمية وذلك بغرض تسهيل تفسيرها. وهذه الاساليب يمكن ان تكون رقميا وحسابيا ويمكن وضعها في جداول مناسبة او بيانيا بحيث يسهل فهمها من خلال الاشكال والرسوم البيانية. و **الاحصاء الاستدلالي** يحتوي تلك الاساليب (الطرق) والتي من خلالها نتمكن من اتخاذ القرارات حول المجتمع الاحصائي وذلك من واقع العينة المسحوبة من هذا المجتمع. وهذه القرارات يتم اتخاذها تحت شروط احتمالية . وتسمى وصف العينة وخصائصها بإحصائية العينة [sample statistics] بينما الخصائص التي تصف المجتمع تسمى م [Parameters] والهدف الاساسي من الاحصاء الاستدلالي (أو التحليلي) هو استنتاج خصائص المجتمع من خصائص عينة سحبت منه . فعند استخدام احصائية العينة [statistics] للاستدلال عن المجتمع الاحصائي ،لا نه لا تتوفر لدينا كل الحقائق عن المجتمع ، فنلجأ للبحث عن طريقة عملية نستطيع من خلالها الوصول الى حقائق عن المجتمع بدرجة من الموثوقية . وهذه الحقائق هي كمية لمعالم المجتمع [parameters] من خلال بيانات العينة المسحوبة من المجتمع عشوائيا .
وينقسم الاستدلال الاحصائي الى قسمين :

1. التقدير [Estimation]
2. اختبارات الفروض [Hypothesis Testing]

التقدير : هو تقدير معالم المجتمع الاحصائي ، و التي تكون في الغالب مجهولة ، ويراد الحصول على تقديرات لها من بيانات العينة (فمثلا تقدير متوسط دخل الفرد في بلد معين من بيانات عينة لمداخيل افراد من المجتمع ، او تقدير نمط الاستهلاك لمنتجات اللحوم في مدينة معينة من عينة من سكان هذه المدينة ونمط استهلاكهم أو تقدير مستوى التحصيل العلمي لطلاب الجامعة من عينة من الطلاب ينتمون الى هذه الجامعة ..الخ) . و توجد طريقتان للتقدير هما : التقدير بنقطة [Point estimate] والتقدير بفترة [Interval estimate]

- **التقدير بنقطة** : هو عبارة عن قيمة واحدة (عدد) يستخدم لتقدير المعلمة المجهولة للمجتمع الاحصائي . وهذا التقدير عبارة عن نقطة (عدد) يسمى التقدير النقطي .
- **التقدير بفترة** : في تقدير الفترة نحصل على مدى (Range) تحدد بحدين

(الحد الاعلى والحد الادنى) نحصل عليهما من العينة المسحوبة من المجتمع . وتقدير الفترة يحوى على اكثر من قيمة يصف فترة من القيم أو أطار يمكن ان تقع بينهما معالم المجتمع .

9.1 : خصائص التقدير الجيد

عدم التحيز [Unbiased]

أن الإحصائية تكون مقدراً غير متحيز للمعلمة عندما يكون متوسط توزيع المعاينة الإحصائية يساوي معلمة المجتمع المقابلة وألا فيسمى مقدراً متحيزا . إن تقدير الوسط الحسابي $\bar{X}$ للعينة المسحوبة عشوائياً من مجتمع مساوياً μ الوسط الحسابي للمجتمع عندها يكون التقدير غير متحيز أي أن:

$$\bar{X} = \mu$$

$$E(\bar{X}) - \mu = 0$$

$$E(\bar{X}) = \mu$$

نقول أن X مقدر نقطة غير متحيز لمتوسط المجتمع μ.

- **التوافق [Consistency]**

يقال عن العينة مقدراً متوافقاً لمعلمة المجتمع المجهولة حال اقتراب متوسط توزيع معاينته من المعلمة المجهولة واقتراب تباين توزيع معاينته من الصفر كلما ازداد حجم العينة . إن المتوسط الحسابي $\bar{X}$ مقدراً متسقاً لمتوسط المجتمع μ كلما يزداد اقتراب تباينه من الصفر و كلما زاد حجم العينة n وبالتالي يزداد الاقتراب من المعلمة نفسها.

- **الكفاءة [Efficiency]**

إذا كان توزيع المعاينة لإحصائيتين متساويتا الوسط الحسابي فالإحصائية ذات التباين الأقل تسمى مقدر كفوء والأخرى تسمى مقدر غير كفوء والقيمة المقابلة للإحصائية تسمى تقدير كفوء أو غير كفوء على الترتيب وكلاهما تقديرين غير متحيزين للمعلمة، كما أن ولجميع الإحصائيات التي توزيع المعاينة لها له نفس الوسط الحسابي فالإحصائية ذات التباين الأقل يسمى أحيانا التقدير الأكثر كفاءة.

- **الكفاية [sufficiency]**

نقول للمقدار كاف اذا استخدم معلومات إضافية من العينة لا يتصف بها مقدر اخر وذلك لاستخلاص معلومات عن المجتمع المراد تقدير معالمه . وهذه الخاصية تشير الى المقدرة التي يتمتع بها الاحصائيون في التعامل مع البيانات .

- **تقدير النقطة لمعالم المجتمع**

غالبا في المجتمع الاحصائي ، لا نعرف معالمه مثل الوسط الحسابي [μ] وتباينه [σ^2]. لا جل ذلك نلجأ الى أخذ عينة من المجتمع . ونوجد الوسط الحسابي للعينة [$\bar{x}$] . ونعتبره تقديرا للوسط الحسابي للمجتمع . ونوجد تباين العينة [S^2] ونعتبره تقديرا لتباين المجتمع وتباينه [σ^2]. في كلتا الحالتين نكون قد اوجدنا تقديرا لنقطة لمعلمة من معالم المجتمع.

- **تقدير الفترة [Interval Estimate]**

الغرض الاساسي من جمع البيانات بأسلوب المعاينة [Sampling] هو دراسة خصاص المجتمع وهذا يمكن فقط من خلال تقدير النقطة او تقدير الفترة . وتقدير الفترة يصف فترة من القيم والتي يمكن ان تقع ضمنه معالم المجتمع . و تقدير النقطة (يعطي قيمة واحدة للمعلمة) بينما تقدير الفترة (يعطي قيمتين والذي يقدر ان تقع بينهما المعلمة). ويحدد تقدير الفترة بقيمتين ، القيمة العليا (الحد الاعلى) والقيمة الدنيا (الحد الأدنى) .

أن تقدير الفترة يشير الى تقدير المعلمة لفترة عشوائية، تسمى فترة الثقة[confidence Interval]

وبمقارنة الطريقتين في التقدير نجد ان تقدير النقطة له الافضلية حيث أنه يعطي قيمة حقيقية للمعلمة . وهذه الميزة يمكن ان تكون ايضا من عيوب هذه الطريقة . حيث ان النقطة المنفردة على خط الاعداد الحقيقية لا توضح لنا مدى قرب القيمة المقدرة الى المعلمة المراد تقديرها . وايضا في الدراسات والابحاث العلمية لا يحبذ ان تكون قيمة واحدة مؤكدة للمعلمة ويفضل ان تكون هناك درجات ثقة وتكون القيمة المقدرة ضمن اطار او فترة محددة تسمى فترة الثقة .

9.2: تقدير الوسط الحسابي لمجتمع معلوم التباين

<u>**نظرية**</u> : تنص نظرية النهاية المركزية [Central Limit Theorem] ،بأنه إذا كان لدينا مجتمع وسطه الحسابي μ وتباينه σ^2 و سحبنا منه كل العينات العشوائية المتاحة ، فتوزيع المعاينة للوسط الحسابي $\bar{x}$ سيكون توزيعا قريبا من التوزيع الطبيعي بوسط قدره μ وتباين قدره σ^2/n . ويمكن صياغة ذلك رياضيا كالاتي :

$$X \sim N(\mu ، \sigma^2)$$

$$X \sim N(\mu ، \sigma^2/n)$$

- **الاثبات**

نفترض ان لدينا مجتمع (N) وسحبت منه عدد من العينات من نفس الحجم (n) ، فأن الاوساط الحسابية لهذه العينات هي :

$\bar{x}_1 ، \bar{x}_2 ، \bar{x}_3 ، \ldots\ldots\ldots\ldots ، \bar{x}_n$

و وسط الاوساط يعطى كالاتي :

$$\mu_{\bar{x}} = \frac{\bar{x}_1 ، \bar{x}_2 ، \bar{x}_3 ، \ldots\ldots\ldots\ldots ،}{n}$$

$$\mu_{\bar{x}} = \frac{\sum \bar{x}}{n} = \mu$$

حيث أن :

$\mu_{\bar{x}}$ = وسط الاوساط الحسابية

n = عدد العينات

μ = الوسط الحسابي للمجتمع

x̄ = الوسط الحسابي للعينة

واذا سحبنا عينات عددها n من مجتمع حجمه (N) فأن الوسط الحسابي لهذه العينات ينتمي الى هذا المجتمع ويساوي μ وهذا هو المطلوب اثباته اولا .

كما ان مربع الانحرافات للأوساط الحسابية عن الوسط الحسابي يعطي :

$(\bar{x}_{1-} \ \mu)^2 ، (\bar{x}_{2-} \ \mu)^2 \ (\bar{x}_{2-} \ \mu)^2 \ldots\ldots\ldots\ldots ، (\bar{x}_{n-} \mu)^2$

$$\sigma^2_{\bar{x}} = (\bar{x} - \ \mu)^2 / n$$

حيث أن :

$\sigma^2_{\bar{x}}$ = تباين الوسط الحسابي

$\bar{x}$ = الوسط الحسابي

μ = الوسط الحسابي للمجتمع

n = عدد العينات

$$\sigma^2_{\bar{x}} = \frac{\sum(\bar{x} - \mu)^2}{n} = \frac{\sum\left(\frac{\sum x}{n} - \frac{\sum \bar{x}}{n}\right)}{n} = \frac{\sum\sum(x - \bar{x})^2}{\frac{n^2}{n}} = \frac{\sum\{\sum(x - \bar{x})^2\}}{n^2 . n}$$

$$= \frac{\sum\{n(x - \bar{x})^2\}}{n^2 . n}$$

$$= \frac{\sum(x - \bar{x})^2}{n^2} = \frac{\sum(x - \bar{x})^2}{\frac{n}{n}} = \frac{\sigma^2}{n} ، \qquad \left(\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}\right)$$

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ [*وهو المطلوب الثاني*]

اي ان :

$\bar{X} \sim N(\mu ، \sigma^2/n)$

وهو المطلوب اثباته

قدير الوسط الحسابي لمجتمع معلوم التباين

$\because \bar{X} \sim N(\mu ، \sigma^2/n)$

$\therefore \sigma_{\bar{x}}^2 = \sigma^2/n$

$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ ....(standard error ..SE) الخطأ المعياري

أذا كانت :

$X \sim N(\mu ، \sigma^2)$

فأن :

$Z = \frac{x-\mu}{\sigma} \sim N(0، 1)$

أما بالنسبة لمتوسط العينة

$\bar{X} \sim N(\mu ، \sigma^2/n)$

$Z_{\alpha/2} = = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0،1)$

$Z_{\alpha/2} = = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

بضرب الطرفين في الوسطين

$Z_{\alpha/2} \, . \sigma/\sqrt{n} = \bar{x} - \mu$

$\mu = \bar{x} - Z_{\alpha/2} \, . \sigma/\sqrt{n}$ .... $\bar{x} > \mu$

$\mu = \bar{x} + Z_{\alpha/2} \, . \sigma/\sqrt{n}$ .... $\bar{x} < \mu$

ومن المعادلتين اعلاه :

$$\bar{x} - Z_{\alpha/2} \, . \sigma/\sqrt{n} < \mu < \bar{x} + Z_{\alpha/2} \, . \sigma/\sqrt{n}$$

هذا هو قانون تقدير الفترة للوسط الحسابي معلوم التباين

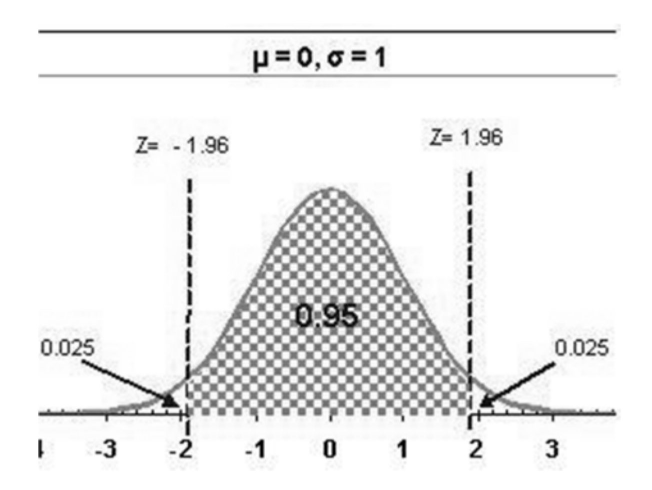
تحديد [$Z_{\alpha/2}$]

$Z_{\alpha/2}$ هي القيمة المعيارية للوسط الحسابي للعينة . وتحدد قيمتها بمقدار مستوى الثقة التي نضعها في التقدير الاحصائي . ومستوى الثقة هو مقدار الاحتمال .فإذا كانت ثقتنا في التقدير 95% فإن الاحتمال :

$P = 1 - \alpha = 0.95$

$\alpha = 1 - 0.95 = 0.05$

حيث إن [α] هي مستوى المعنوية



.

حيث أن المنحنى متماثل فإن المساحة على جانبي المنحنى متساوي

ولتحديد [$Z_{\alpha/2}$ = $Z_{0.025}$]

$0.5 - 0.025 = 0.475$

ومن جدول التوزيع الطبيعي المعياري ، جدول Z، نجد أن $Z_{0.025}$ هي المقابلة للمساحة 0.475 اي أن :

$Z_{0.025} = 0.475 = 1.96$

وذلك باحتمال 95%

$$\bar{x} - \mathbf{1.95} \, .\sigma/\sqrt{n} < \mu < \bar{x} + \mathbf{1.96} \, .\sigma/\sqrt{n}$$

هذا القانون يستخدم لا يجاد فترة الثقة باحتمال 95% للوسط الحسابي للمجتمع معلوم التباين .

ويمكن كتابة القانون اعلاه بصيغة رياضية افضل :

$$\mu = \bar{x} \pm 1.96 \;.\sigma/\sqrt{n}$$

حيث أن :

الحد الاعلى لفترة الثقة = $\bar{x} + 1.96 \;.\; \sigma/\sqrt{n}$

الحد الادنى لفترة الثقة = $\bar{x} - 1.96 \;.\; \sigma/\sqrt{n}$

واذا كان المجتمع طبيعي وسطه الحسابي **μ** وتباينه σ^2 مجهولان . وكانت $\bar{x}$ الوسط الحسابي للعينة و S^2 تباينها و **n** حجم العينة ($n>30$) . و فترة الثقة للوسط الحسابي تعطى بالقانون الاتي :

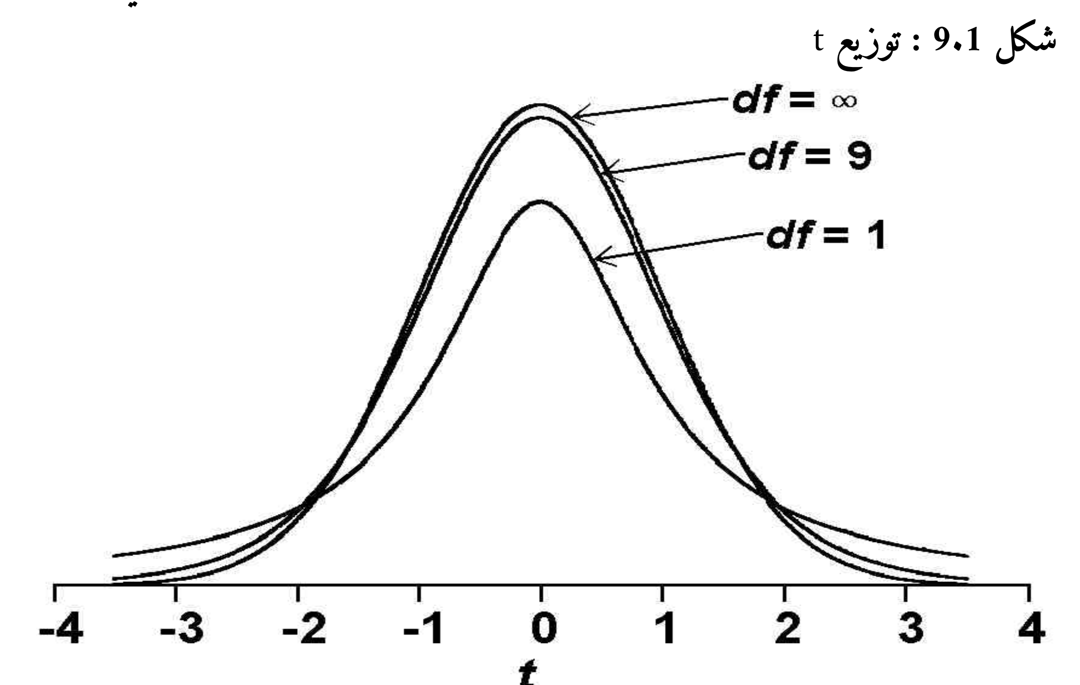
$$\bar{x} - Z_{\alpha/2} \;.s/\sqrt{n} < \mu < \bar{x} + Z_{\alpha/2} \;.s/\sqrt{n}$$

أي أن :

$$\mu = \bar{x} \pm 1.96 \;.s/\sqrt{n}$$

9.3 : فترات الثقة للوسط الحسابي باستخدام توزيع [*t*] : عندما يكون التوزيع طبيعيا ولكن التباين σ^2 مجهول و $n< 30$ فإننا نستخدم توزيع **t** . هذا التوزيع متماثل حول المتوسط . ولكنه اكثر تفلطحا بالمقارنة مع التوزيع الطبيعي المعياري . ولهذا جزء كبير من مساحته تقع عند الاطراف . بينما يوجد توزيع طبيعي معياري واحد ، فان هناك توزيعا **t** مختلفا لكل حجم من العينة وبدرجات حرية مختلفة .. لكن كلما زاد حجم العينة فإن توزيع **t** يقترب من التوزيع الطبيعي المعياري عندما تكون $n \geq 30$ ،وعندئذ فإن توزيع **t** يقترب من التوزيع الطبيعي . حتى يتساويا .

شكل 9.1 : توزيع t



وتعطى قيم t في جدول توزيع t تحت درجات حرية مختلفة . ودرجات الحرية df تعطى بالعلاقة :

$$df = n-1$$

درجات الحرية هي حجم العينة ناقصا واحد . واذا كنا نريد تقدير معلمة واحدة μ ، وفترة الثقة للوسط الحسابي للمجتمع المجهول باستخدام توزيع t تعطى بالعلاقة

$$\bar{x} - t_{\alpha/2} . s/\sqrt{n} < \mu < \bar{x} + t_{\alpha/2} . s/\sqrt{n}$$

أي أن :

$$\mu = \bar{x} \pm t_{\alpha/2} . s/\sqrt{n}$$

- **أمثلة على المتوسطات**

مثال (9.1) : إذا كان الانحراف المعياري لإنتاج اشجار البرتقال في مزرعة يساوي 7 كغ . ،اختيرت عينة عشوائية كبيرة من 36 شجرة و وجد أن متوسط أنتاج الشجرة 20 كغ . فما هي فترة الثقة لمتوسط المجتمع عند مستوى الثقة 90% .

الحل : فترة الثقة لمتوسط المجتمع معلوم التباين يعطى بالقانون الاتي :

$$\mu = \bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$$

$$z_{\alpha/2} = Z0.05 = 1.64$$

$$\mu = \bar{x} \pm 1.64 \cdot \sigma/\sqrt{n}$$

$$= 20 \pm 1.64 \cdot 7/\sqrt{36}$$

$$= 20 \pm (1.64 \times 1.167)$$

$$= 20 \pm 1.913 = 21.913$$

وبذلك فإن فترة الثقة لمتوسط المجتمع لإنتاج اشجار البرتقال عند مستوى ثقة 90% تساوي (18.09 -21.913) . أي أننا واثقون بنسبة 90% بأن متوسط الانتاج لشجرة البرتقال يقع ضمن الفترة [18.09 – 21.91] .(كغ)

مثال (9.2) اختيرت عينة عشوائية من 64 عبوة لمسحوق الصابون فإذا تبين أن متوسط وزن العبوة في العينة هو 2.95 كغ . حسب الانحراف المعياري من العينة ووجد أنه يساوي 0.3 كغ . فما هي حدود الثقة للمتوسط الحقيقي للمجتمع عند مستوى ثقة 95 %

الحل :

حيث إن العينة كبيرة ≤ 30

عند مستوى ثقة 95% (مستوى المعنوية $\alpha = 0.05$)

$$z_{\alpha/2} = z_{0.025} = 1.96$$

$$\mu = \bar{x} \pm 1.96 \cdot \sigma_{\bar{x}}$$

$$\mu = \bar{x} \pm 1.96 \cdot s/\sqrt{n}$$

$$= 2.95 \pm 1.96 \cdot 0.3/\sqrt{64}$$

$$= 2.95 \pm (1.96 \times 0.0375)$$

$$\mu = 2.95 \pm 0.0735$$

أي أننا واثقون بنسبة **95**% بأن متوسط ،وزن عبوة الصابون تقع في الفترة (2.87 – 3.02 كغ)

مثال (9.3) : عينة عشوائية من 25 مفردة بمتوسط 80 ، أخذت من مجتمع توزيعه طبيعي ، بانحراف معياري 30 . أوجد فترات الثقة الاتية لوسط المجتمع غير المعلوم.

I. 90%

II. 95%

III. 99%

IV. فسر النتائج

الحل : حيث أن σ معلومة ، فأننا نستخدم التوزيع الطبيعي المعياري .

- **عند مستوى ثقة 90%**

$$\mu = \bar{x} \pm z_{\alpha/2} \cdot \sigma_{\bar{x}}$$
$$\mu = \bar{x} \pm 1.64 \cdot \sigma_{\bar{x}}$$
$$\mu = \bar{x} \pm 1.64 \cdot \sigma/\sqrt{n}$$
$$= 80 \pm 1.64 \cdot 30/\sqrt{25}$$
$$= 80 \pm 1.64 \times 6$$
$$= 80 \pm 9.84$$

أي أن الوسط الحسابي للمجتمع μ يقع بين 70،16 و 89،84 وذلك عند مستوى ثقة 90%

- عند مستوى ثقة 95%

$$\mu = \bar{x} \pm 1.96 \cdot \sigma/\sqrt{n}$$
$$= \bar{x} \pm 1.96 \cdot 30/\sqrt{25}$$
$$= 80 \pm 1.96 \times 6$$
$$= 80 \pm 1.96 \times 6$$
$$= 80 \pm 11.76$$

أي أن الوسط الحسابي للمجتمع μ يقع في الفترة 68،24 - 91،76 وذلك عند مستوى ثقة 95 %

- عند مستوى ثقة 99 %

$$\mu = \bar{x} \pm 2.58 \cdot \sigma/\sqrt{n}$$
$$= 80 \pm 2.58 \cdot \left(\frac{30}{\sqrt{25}}\right)$$
$$= 80 \pm 2.58 \times 6$$
$$= 80 \pm 15.48$$

أي أن الوسط الحسابي للمجتمع μ يقع في الفترة (64.52 - 95.48) وذلك عند مستوى ثقة 99%

- النتائج أعلاه ، تشير الى أنه مع زيادة درجة الثقة فأن حجم فترة الثقة يزداد أيضا (اي أن هناك علاقة طردية بين درجة الثقة وفترة الثقة) ويصبح التقدير اقل دقة . **ولكن** درجة الثقة المرتبطة بفترة ثقة ضيقة جيدا قد تكون منخفضة بدرجة تفقد معها المعنى . وجرت العادة كتقليد بأن فترات الثقة الاكثر استخداما هي 90% ، 95% و 99%

مثال (9.4) اختيرت عينة عشوائية من 25 عبوة مسحوق تنظيف ووجد أن متوسط وزن العبوة في العينة هو 2.95 كغ وأحتسب الانحراف المعياري من العينة ووجد أنه 0.30 . فما هي حدود الثقة للمتوسط الحقيقي لوزن العبوة عند مستوى ثقة 95% على افتراض أن توزيع عبوات مسحوق التنظيف طبيعيا .

الحل

بما أن العينة صغيرة < 30 وعند مستوى ثقة 95% ، فأن مستوى المعنوية $\alpha = 0.05$

$$t_{0.025} = t_{\alpha/2}$$

عند درجة حرية : $df = n-1 = 25-1 = 24$

قيمة $t_{\alpha/2}$ = 2.064 (من جدول توزيع t)

$$\mu = \bar{x} \pm t_{\alpha/2} \cdot s/\sqrt{n}$$

$$= \bar{x} \pm 2.064 \cdot 0.30/\sqrt{25}$$

$$\mu = 2.95 \pm 0.124$$

إن فترة الثقة هي (**2.83 – 3.074**) عند مستوى ثقة 95%

أي أن : $3.074 > \mu > 2.83$

مثال (9.5) سحبت عينة عشوائية مكونة من 9 مصابيح كهربائية بمتوسط عمر 300 ساعة احتراق وانحراف معياري 45 ساعة من شحنة كهربائية كبيرة من المصابيح معروف ان عمر تشغيلها يتبع التوزيع الطبيعي . أوجد فترة الثقة 90% لوسط عمر التشغيل غير المعلوم .

الحل : عند درجة حرية : $df = n-1 = 9-1=8$

قيمة $t_{\alpha/2} = t_{0.05} = 1.860$ (من جدول توزيع t)

$$\mu = \bar{x} \pm t_{\alpha/2} \cdot s/\sqrt{n}$$

$= 300 \pm 1.860 \times 45/\sqrt{9}$

$\mu = 300 \pm 27.9$

أن فترة الثقة هي (272 - 328) عند مستوى ثقة 90%

أي أن : $328 > \mu > 272$ *عند مستوى ثقة 90%*

مثال (9،6) يرغب مدير مصنع في تقدير متوسط عدد الدقائق التي يأخذها العامل لإنجاز عملية صناعية معينة في حدود (3 ±) د قائق وبدرجة ثقة 90% . ويعلم المدير من خبرته الماضية أن الانحراف المعياري هو 15 دقيقة . ما هو الحد الادنى للعينة المطلوبة .

الحل :

$Z_{\alpha/2} = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$

$Z_{\alpha/2} . \sigma_{\bar{x}} = \bar{x} - \mu$

$1.64 . \frac{\sigma}{\sqrt{n}} = \bar{x} - \mu$

$1.64 . \frac{15}{\sqrt{n}} = 3$

$3. \sqrt{n} = 1.64 \times 15$

$\sqrt{n} = \frac{1.64 \times 15}{3} = 8.20$

$\sqrt{n} = 8.20$

$n = 67.24 \sim 68$

$n = 68$ .... (حجم العينة المطلوبة)

ملحوظة : المتوسط الحقيقي = متوسط العينة + حجم الخطأ

$\mu = \bar{x} \pm$ حجم الخطأ

مثال (9،7) يراد تقدير متوسط الانفاق الشهري من الوقود في حدود ± 1.495 والانحراف المعياري 10. ما هو حجم العينة اللازم عند مستوى ثقة 95% .

الحل :

$\bar{x} - \mu = Z_{\alpha/2} . \sigma_{\bar{x}}$

$\bar{x} - \mu = Z_{\alpha/2} . \frac{\sigma}{\sqrt{n}}$

$1.495 = 1.96 . \frac{10}{\sqrt{n}}$

$\sqrt{n} . 1.495 = 19.6$

$\sqrt{n} = \frac{19.6}{1.495}$

$$n = [\ 19.6/1.495\ ]^2$$

$$n = 171.9 = 172$$

مثال (9.8) إذا كان الانحراف المعياري لإنتاج اشجار البرتقال في مزرعة كبيرة يساوي 7 كغ .، اختيرت عينة عشوائية من 36 شجرة للتعرف على متوسط الانتاجية للشجرة . فما هو احتمال أن يكون المتوسط في حدود (2 $\pm$) من المتوسط الحقيقي للإنتاجية

الحل:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7}{\sqrt{36}} = 1.167$$

$$Z = = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

$$Z = = \frac{(\mu - 2) - \mu}{1.67} \quad \ldots\ldots\ldots \quad (\bar{x} = \mu - 2)$$

$$Z = = \frac{(\mu - 2) - \mu}{1.67} = \frac{-2}{1.67} = -1.71$$

والمساحة المقابلة لقيمة **(z= -1.71)** في جدول التوزيع الطبيعي المعياري هي **z=0.4564**

$$Z = = \frac{(\mu + 2) - \mu}{1.67} \quad \ldots\ldots\ldots \quad (\bar{x} = \mu + 2)$$

$$Z = = \frac{(\mu + 2) - \mu}{1.67} = \frac{+2}{1.67} = +1.71$$

والمساحة المقابلة لقيمة **(z= +1.71)** في جدول التوزيع الطبيعي المعياري هي **z=0.4564**

مجموع المساحتين = 0.4564 + 0.4565 = 0.913

(0.913) يمثل احتمال أن يكون المتوسط في حدود (2 $\pm$) من المتوسط الحقيقي وهو متوسط إنتاجية شجرة البرتقال.

مثال (9.9) أذكر اي توزيع ينبغي استخدامه لا يجاد فترات الثقة للوسط غير المعلوم للمجتمع من عينة عشوائية مأخوذة من المجتمع في الحالتين التاليتين :

- **n = 64 ، s = 8**
- **n =20 ، s = 10 (المجتمع لا يتبع التوزيع الطبيعي)**
- {يترك الحل للطالب }

- **معامل التصحيح**

لعينات عشوائية ذات الحجم n مأخوذة من مجتمع له متوسط μ وانحراف معياري σ فأن $\bar{x}$ تكون مساوية لمتوسط المجتمع μ والخطأ المعياري يعطى بالقانون ألاتي :

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \ldots(1) \qquad \text{or} \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \ldots(2)$$

حيث تستخدم معادلة (2) للمجتمعات المحدودة ذات الحجم **N** عندما تكون $n \geq 0.05$

ويسمى المقدار $\sqrt{\frac{N-n}{N-1}} \ldots$ **معامل التصحيح** و $\sigma_{\bar{X}}$ يسمى الخطأ المعياري

مثال(9.10) افترض أن المجتمع يتكون من 1000 عنصر بوسط حسابي **20** وحدة وانحراف معياري 14 وحدة . أوجد الوسط الحسابي والانحراف المعياري لعينة حجمها **36** .

الحل

$$\mu_{\bar{X}} = \mu = 20$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = 2$$

لو كانت **n** تساوى **64** بدلا من **36** (بحيث **n > 0.05 N**) ، فإن

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{14}{\sqrt{64}}\sqrt{\frac{1000-64}{1000-1}} = \frac{14}{8}\sqrt{\frac{936}{999}} = (1.75)(0.9679) = 1.69$$

بدلا من 1.75 = $\sigma_{\bar{X}}$ بدون **معامل التصحيح للمجتمعات المحدودة .**

مثال (9.11) جمعية خدمة المجتمع تريد تقدير الدخل السنوي (700) أسرة تقطن احدى الاحياء الشعبية المقسمة الى اربعة قطاعات . أخذت عينة عشوائية بسيطة حجمها 50 بوسط حسابي 5000 $ و انحراف معياري $s = 950$ $. الجمعية طلبت من مندوبها تقدير الفترة للدخل السنوي للأسرة (عددها 700) بحيث انها واثقة بدرجة ثقة 90% بأن متوسط دخل المجتمع يقع في تلك الفترة .

الحل :

بما أن :

$$n > 0.05$$

إذا :

$$\bar{x} - \mu = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$\sigma_{\bar{x}} = \frac{950}{\sqrt{50}} \cdot \sqrt{\frac{700-50}{700-1}}$

$\sigma_{\bar{x}} = \frac{950}{7.07} \cdot \sqrt{\frac{650}{699}} = (134.37)(0.9643) = 129.57$

$\sigma_{\bar{x}} = 129.57$ ... ... الخطأ المعياري

$\mu = \bar{x} \pm Z_{\alpha/2} \cdot \sigma_{\bar{x}}$

$\mu = \bar{x} \pm (1.64)(129.57)$

$\mu = 5000 \pm (1.64)(129.57)$

$\mu = 5000 \pm 212.50$

$\mu = 5212.5$ .... **الحد الاعلى للفترة**

$\mu = 5000 - 212.50$

.... **الحد الادنى للفترة**

$\mu = \{ 4778.5 - 5012.5 \}$

بدرجة ثقة 90% فأن متوسط الدخل السنوي (700) أسرة يقع بين 4778.5 $ و 5012.5 **$.**

9.4 : تقدير فترة الثقة للفرق بين متوسطين

إذا كان لدينا مجتمعين مستقلين كل منهما بتوزيع احتمالي معين ، متوسط المجتمع الاول ، μ_1 وتباينه σ_1^2 ، ومتوسط المجتمع الثاني μ_2 وتباينه σ_2^2 ، واذا كان اهتمامنا يتركز على الفرق بين متوسطي المجمعين($\mu_1 - \mu_2$) ، فإننا نختار عينتين عشوائيتين مستقلتين n_1 من المجتمع الاول وأخرى n_2 من المجتمع الثاني ونحسب كلا من المتوسط الحسابي والتباين للعينة الاولى $\bar{x}_1$ ، s_1^2 ، ثم نحسب كلا من المتوسط الحسابي والتباين للعينة الثاني $\bar{x}_2$ ، s_2^2 .

نحسب فترة الثقة للفرق بين المتوسطين عند مستوى ثقه معين وفق المعادلة الأتية :

$$\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 \pm z_{\sigma/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

حيث يمثل المقدار ($\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$) **انخطأ المعياري للفرق بين متوسطين .** وإذا لم تكن قيم التباين معلومة للمجتمعين ، نستخدم قيم التباين s_2^2، s_1^2 المحسوبة من العينات حيث تعتبر تقديرا جيدا اذا كان حجم العينات كبيرة ($n \geq 30$) .

مثال (9.12) اختيرت عينة عشوائية من المجتمع الاول بحجم 144 مفردة ، ووجد أن المتوسط يساوي 65 والتباين 25 . واختيرت من المجتمع الثاني مستقلة عن الأول بحجم 100 و وجد ان المتوسط يساوي 60 والتباين 16 . فإذا كانت قيم التباين غير معروفة للمجتمعين وكانت العينات مستقلة . أحسب فترة الثقة للفرق بين متوسطي المجتمعين وذلك عند مستوى ثقة 95% .

الحل :

بما أن :

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z_{\sigma/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\mu_1 - \mu_2 = (60 - 65) \pm 1.96 \sqrt{\frac{25}{144} + \frac{16}{100}}$$

$$= 5 \pm 1.96\ (0.5775)$$

$$= 5 \pm 1.1319$$

أي أن فترة الثقة للفرق بين متوسطي المجتمعين عند مستوى ثقة 95% هي :

$$3.8681 \leq \mu 1 - \mu 2 \leq 6.1319$$

- **حدود الثقة للفرق بين متوسطين / عينات صغيرة**

يتطلب تقدير فترة الثقة للفرق بين متوسطين لعينات صغيرة (< 30) توفر الشروط التالية :

I. توزيع مجتمعي الدراسة طبيعي (n_1، n_2)

II. العينات مستقلة

III. التباين أو الانحراف المعياري غير معروف للمجتمعين ولكنه (يفترض) متساو للمجتمعين ($\sigma_2^2 = \sigma_1^2$) ولذلك يستخدم التباين من العينتين كتقدير للمجتمعين

نحسب فترة الثقة للفرق بين المتوسطين عند مستوى ثقة ($1-\alpha$) بالمعادلة الأتية :

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm t_{\sigma/2} \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}$$

$$\text{الخطأ المعياري للفرق بين متوسطين} = \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}$$

وتحدد قيمة $t_{\sigma/2}$ عند درجة حرية : $df = n_1 + n_2$

قيمة s^2 التباين التجميعي تحتسب بالمعادلة التالية :

$$s^2 = \frac{s_1^2\,(n_1 - 1) + s_2^2\,(n_2 - 1)}{n_1 + n_2 - 2}$$

مثال (**9.13**) قامت احدى المؤسسات الصحية بدراسة للمقارنة بين متوسط محتوى البروتين في الغذاء اليومي لمجموعتين من السكان . فإذا كانت قيم التباين غير معلوم للمجتمعين ولكنه يفترض أنهما متساوية . وتوزيع المجتمعين طبيعي . اختيرت عينات عشوائية مستقلة تتكون من 10 فرد من المجتمع الاول و 15 فردا من المجتمع الثاني وتم احتساب كمية البروتين في غذائهم و وجد من العينات ان متوسط المحتوى البروتيني للعينة من المجتمع الاول يساوي 88 غرام والتباين يساوي 11 غرام . ومتوسط المحتوى البروتيني للعينة من المجتمع الثاني يساوي 77 غرام والتباين يساوي 9 غرام . أحسب فترة الثقة بين متوسطي المجتمعين عند مستوى ثقة 95%.

الحل:

نحسب أولا التباين التجميعي باستخدام المعادلة :

$$s^2 = \frac{11(10-1) + 9(15-1)}{10+15-2} = \frac{11(9) + 9(14)}{23} = 9.783$$

و تحدد قيمة $t_{\sigma/2}$ وهي $t_{0.025}$ عند درجة حرية :

$$df = n_1 + n_2 = 10+15-2 = 23$$

وتساوي هذه القيمة $t_{0.05} = 1.714$

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}$$

$$= (88 - 77) \pm 1.714\sqrt{\frac{9.783}{10} + \frac{9.783}{15}}$$

$$= 11 \pm 2.1886$$

$$= (8.8114 - 13.1886)$$

$$13.1886 \geq \mu_1 - \mu_2 \geq 8.8114$$

9.5 : حدود الثقة للنسبة لمجتمع من عينة كبيرة ($n \geq 30$)

- **حدود الثقة لنسبة واحدة**

نحتاج لمعرفة توزيع المعاينة للمتوسطات حتى يمكن الحصول على استنتاجات حول المتوسط في المجتمع وبالمثل نحتاج لمعرة توزيع المعاينة للنسب ، أي التوزيع الاحتمالي للمتغير العشوائي ($\bar{P}$) حتى نتمكن من الحصول على استنتاجات حول النسبة في المجتمع (P) وذلك باستخدام المساحة تحت المنحنى الطبيعي المعياري . و لا جل ذلك من الضرورة اولا الحصول على النسبة من العينة والخطأ المعياري للنسبة . ونستخدم اختبار التوزيع الطبيعي المعياري Z في حالة تقدير الفترة للنسبة ولا نستخدم اختبار t

- خطوات أيجاد حدود الثقة لنسبة المجتمع

I. لمستوى ثقة ($1-\alpha$) نستخدم جدول قيم Z لا يجاد قيمة $Z_{\alpha/2}$

II. حدود الثقة للنسبة في المجتمع (P) يعطى بالقانون :

$$P = \bar{P} \pm Z\,\sigma_p$$

$$P = \bar{P} \pm Z\sqrt{\frac{p(1-p)}{n}}$$

$$P = \bar{P} \pm Z\sqrt{\frac{p(1-p)}{n}}\cdot\sqrt{\frac{N-n}{N-1}} \quad \text{..............} \; n > 0.05\,N$$

9.6 : فترة الثقة للفرق بين نسبتين ($P_1 - P_2$) هي :

$$(P_1 - P_2) = (\bar{P}_1 - \bar{P}_2) \pm Z_{\alpha/2}\sqrt{\frac{\bar{P}1(1-\bar{P}1)}{n1} + \frac{\bar{P}2\,(1-\bar{P}2)}{n2}}$$

مثال (9.14) أخذت عينة من 100 أسرة ووجد ان 34 من بينهم يملكون جهازا للهاتف . أوجد حدود الثقة للنسبة للمجتمع عند مستوى ثقة 95%.

الحل

لمستوى ثقة ($1-\alpha$) فأن مستوى المعنوية هو 5% . ونستخدم جدول المنحنى الطبيعي المعياري Z ، لا يجاد المساحة تحت المنحنى للقيمة المعيارية $Z_{\alpha/2}$

$$Z_{\alpha/2} = Z_{0.025} \pm 1.96$$

فترة الثقة للنسبة في المجتمع (P)

نسبة المالكين للهاتف = $\overline{P}$

$$\overline{P} = 34/100 = 0.34$$

نحسب الخطأ المعياري للنسبة باستخدام النسبة من العينة لان النسبة للمجتمع غير موجود

$$P = \overline{P} \pm Z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$P = 0.34 \pm 1.96 \sqrt{\frac{(0.34)(0.66)}{100}}$$

$$P = 0.34 \pm 1.96\ (0.047)$$

$$P = 0.34 \pm (0.09)$$

$$P = \{0.43 - 0.25\}$$

أي أننا على ثقة بنسبة 95% ، أن النسبة الحقيقية تقع بين القيمتين (0.25 - 0.43)

مثال (9.15) في عينة عشوائية حجمها 100 عامل من مصنع به 1200 عامل ، وجد ان 60 يفضلون الاشتراك ي نظام المعاشات كأفراد بدلا من الاشتراك في مشروع معاشات خاص بالشركة . أوجد فترة الثقة 95% لنسبة العاملين الذين يفضلون مشروعات معاشات فردية .

الحل

$$\overline{P} = 60/100 = 0.60$$

$$P = \overline{P} \pm Z_{\alpha/2}\, \sigma_p$$

حيث أن $N \geq 0.05$

$$P = \overline{P} \pm Z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \quad \ldots\ldots\ldots\ldots\ldots\ n > 0.05\ N$$

$$P = 0.60 \pm 1.96 \sqrt{\frac{(0.60)(0.40)}{100}} \cdot \sqrt{\frac{1200-100}{1200-1}}$$

$$P = 0.60 \pm 1.96\ (0.048)\ (0.96)$$

$$= 0.60 \pm 0.0903$$

وعليه فأن P (نسبة العاملين الذين يفضلون معاشات فردية) تقع في الفترة 0.61 و 0.69 عند مستوى ثقة 95% .

مثال (16) أخذت عينة عشوائية حجمها 80 طالبا من طلبة الثانوية العامة القسم العلمي ووجد أن 22 منهم يدخنون . واخذت عينة عشوائية اخرى حجمها 60 طالبا من طلبة الثانوية العامة القسم الادبي ووجد ان 25 منهم يدخنون .أوجد فترة الثقة 95% للفرق بين النسبتين .

الحل :

نفرض أن نسبة الطلاب القسم الادبي المدخنين = P_1= 25 /60 = 0.416

نفرض أن نسبة الطلاب القسم العلمي المدخنين = P_2= 22/80= 0.275

فترة الثقة للفرق بين نسبتين عند مستوى ثقة 95% هي:

$$(P_1 - P_2) = (\bar{P}_1 - \bar{P}_2) \pm Z_{\alpha/2}\sqrt{\frac{\bar{P}1(1-\bar{P}1)}{n1} + \frac{\bar{P}2\,(1-\bar{P}2)}{n2}}$$

$$(P_1 - P_2) = (0.416 - 0.275) \pm 1.96\sqrt{\frac{0.416(0.584)}{80} + \frac{0.275\,(0.725)}{60}}$$

$$= 0.141 \pm 1.96\sqrt{0.0062659}$$

$$= 0.141 \pm 1.96\,(0.079)$$

$$= 0.141 \pm 0.155$$

أذا فترة الثقة تقع بين (-0.014، 0.296) عند مستوى ثقة 95% .

تمارين الفصل التاسع

1. أخذت عينة عشوائية من 25 مفردة بمتوسط 80 وانحراف معياري 30 من مجتمع مكون من 1000 مفردة ويتبع التوزيع الطبيعي . أوجد فترات الثقة الاتية لوسط المجتمع غير المعلوم . (أ) 90% ، (ب) 95% ، (ج) 99% .
2. يرغب صاحب مصنع في تقدير حجم العينة (n) فما هو هذا الحجم حتى يمكنه التأكد من أن تقديره باحتمال 95% لن يكون خاطئا بأكثر من 5 وحدات معيبة . اذا علم ان الانحراف المعياري 20 وحدة .
3. أخذت عينة عشوائية مكونة من 16 مفردة بمتوسط حسابي 50 ،انحراف معياري 10 من مجتمع كبير جدا يتبع التوزيع الطبيعي . أوجد فترة 95% للوسط غير المعلوم للمجتمع .

4. أخذت عينة عشوائية حجمها 144 بوسط حسابي مقداره 100 وانحراف معياري كقداره 60 من مجتمع حجمه 1000 . أوجد فترة الثقة 95% لوسط المجتمع غير المعلوم .
5. لدى بنك محلي صغير حساب ادخار شخصي برصيد متوسط قدره 3000 $ وانحراف معياري 1200$ ز أخذ البنك عينة عشوائية من 100 حساب ، ما احتمال أن متوسط المدخرات لهذه الحسابات المائة سيكون اقل من 2800$.
6. سحبت عينة عشوائية مكونة من 9 مفردات من المصابيح الكهربائية بمتوسط عمر 300 ساعة وانحراف معياري 45 ساعة من شحنة كهربائية معروف ان عمر تشغليها يتبع التوزيع الطبيعي ، أوجد فترة الثقة 90% لوسط عمر التشغيل غير المعلوم .
7. عينة عشوائية من 64 مفردة وسطها 50 وانحرافها المعياري20 أخذت من مجتمع مفرداته 800 ، أوجد فترة الثقة لوسط المجتمع غير المعلوم عند مستوى ثقة 95% .
8. أفترض أن 50% من 60 مصنعا في اقليم (A) تخضع لمعايير مكافحة التلوث بينما 40% فقط من 40 مصنعا في اقليم (B) تخضع لنفس المعايير . هل نسبة المصانع التي تخضع لمعايير مكافحة التلوث أكبر معنويا في اقليم (A) عنها في اقليم (B) عند مستوى معنوية 5% و 10% .
9. اختيرت عينة عشوائية من المجتمع الاول بحجم 100 من مجتمع و وجد أن المتوسط يساوي 65 والتباين 25 . واختيرت من المجتمع الثاني عينة عشوائية مستقلة عن المجتمع الاول بحجم 80 ووجد أن المتوسط يساوي 60 والتباين 16 . فاذا كانت قيم التباين غير معلومة للمجتمعين وكانت العينات مستقلة ، أحسب فترة الثقة للفرق بين متوسطي المجتمعين عند مستوى ثقة 90%.

وجدت أدارة التعليم لا حدى الولايات الامريكية ، أن في عينة من 100 شخص مختارين عشوائيا من بين الملتحقين بالجامعات 40% منهم قد حصلوا على درجات جامعية . أوجد فترة 99% ثقة لنسبة الحاصلين على درجات جامعية من بين جميع الملتحقين بالجامعة .



الفصل العاشر

اختبار الفرضيات

[Hypothesis Testing]





1 : مفهوم اختبار الفرضيات

اختبار الفرضيات عن خصائص المجتمع هو الجانب الثاني من جانبي الاستدلال الاحصائي . ويمثل التقدير الجانب الاول منه والذي تم تناوله بالتفصيل في الفصل التاسع . وكثيرا ما نستخدم الاستدلال الاحصائي (الإحصاء التحليلي) لاتخاذ القرارات حول متوسط أو نسبة المجتمع أو الفروق بين المتوسطات أو النسب لمجتمعين أو أكثر .وذلك للتعرف على خصائص المجتمع من واقع إحصائيات العينة.

قد نحتاج مثلا الى معرفة فيما اذا كان متوسط الاستهلاك الفردي للحوم في مدينة ما قد أنخفض عن 20 كغ عام 2010 مقارنة بعام 2000 . أو هل نسبة الامية للبالغين قد انخفضت عن 40% في نفس المدينة لنفس الفترة .

لذلك نحتاج الى اختبار أو تقييم الادعاءات أو الفرضيات حول قيم معالم المجتمع مثل المتوسطات والنسب والتباين.

ويمثل اختبار الفرضيات أحد الطرق المتبعة لاتخاذ قرار بشأن هذه الادعاءات حول المجتمع الاحصائي من حيث قبولها أو رفضها وتحديد ما اذا كانت النتائج المشاهدة تختلف معنويا عن المعالم المعلومة للمجتمع . وبالتالي فأن اختبار الفرضيات يشكل قاعدة اساسية لاتخاذ القرار .

في اختبار الفرضيات نبدأ بعمل فرضية أو ادعاء عن خاصية المجتمع غير المعلومة. ونأخذ عينة عشوائية ، وعلى أساس الخاصية المناظرة في العينة إما أن نقبل أو أن نرفض الفرضية بدرجة ثقة محددة . والفرضية هي أدعاء حول صحة أو قيمة شيء ما ،اختبار الفرضية هو تقدير مدى صحة هذا الادعاء (الفرضية) .

10.2 : فرضيات اختبار الفروض

يتضمن اختبار الفروض على فرضيتين هما : فرضة العدم والفرضية البديلة .

- **فرضية العدم : [Null Hypothesis]**

فرضية العدم (ويرمز لها بالرمز H_o) هي الفرضية التي نتمسك بها ولا نرفضها الا اذا توفرت دلائل قوية من العينة تقود الى رفضها . وتعني كلمة العدم (Null) أن الادعاء باطل أو فارغ أو عدم وجود فرق بين معلمة المجتمع والقيمة المدعاة .وفرضية العدم هي التي تكون موضع الاختبار وهل سترفض لمصلحة الفرضية البديلة أم لا ترفض (بمعنى أنها تقبل) بناء على الدلائل التي توفرها العينة . وفي حالة اختبار فرضية حول احد معالم المجتمع فأنها تشمل قيمة واحدة

فقط . فمثلا اذا اردنا اختبار فرضية أن متوسط استهلاك الخبز للفرد سنويا في أحدى المدن عام 2002 هو 160 كغ فأن فرضية العدم تصاغ كالاتي :

$$H_0 : \mu = 160$$

- **الفرضية البديلة [Alternative Hypothesis]**

الفرضية البديلة (ويرمز لها بالرمز H_1) . وهي الفرضية التي يضعها الباحث كبديل عن فرضية العدم وتبنى على اساس أن فرضية العدم غير صحيحة . ففي المثال السابق حول استهلاك الخبز فأن الفرضية البديلة هي أن متوسط استهلاك الخبز في المدينة هو أقل من 160 كغ (أو أكثر أو تختلف عنها) .بناء على معلومات تم الحصول عليها من العينة ، وعليه فأن الفرضية البديلة (في المثال السابق) تصاغ كالاتي :

$H_1 : \mu < 160$

$H_1 : \mu > 160$

وعندما نرفض فرضية العدم فأننا نقبل الفرضية البديلة عند مستوى ثقة معينة . وتوجد صيغ أخرى للفرضية البديلة . فمثلا أذا ارادت شركة للمواد الغذائية ، تقوم بإنتاج نوع معين من المواد الغذائية في عبوات 100 جرام .و اذا اردنا ان نتأكد فيما اذا كان وزن العبوات لا يقل او يزيد عن 100 غرام فان الفرضية لهذا الاختبار تصاغ كالاتي:

$H_0 : \mu = 160$

$H_1 : \mu \neq 160$

وفرضية العدم واحدة مهما كانت الرضية البديلة سواء اكانت اقل أو أكثر (أو تختلف)

10.3 : تصنيف الاخطاء في اختبار الفرضيات

في اختبار الفرضيات يمكن أن نرتكب نوعين من الاخطاء :

- **الخطأ من النوع الا ول (Type I error)** ويسمى بالخطأ (α) ويحدث هذا الخطأ عند رفض فرضية العدم وهي في الواقع صحيحة.
- **الخطأ من النوع الثاني (Type II error)** ويسمى بالخطأ (β) ويحدث هذا الخطأ عند قبول فرضية العدم وهي في الواقع خاطئة .

جدول (10.1) **تصنيف أخطاء اختبار الفروض**

القرار	فرض العدم H_0	
رفض (H_0)	**صحيحة**	**خاطئة**

	الخطأ α	قرار صحيح
قبول (H_0)	قرار صحيح	الخطأ ß

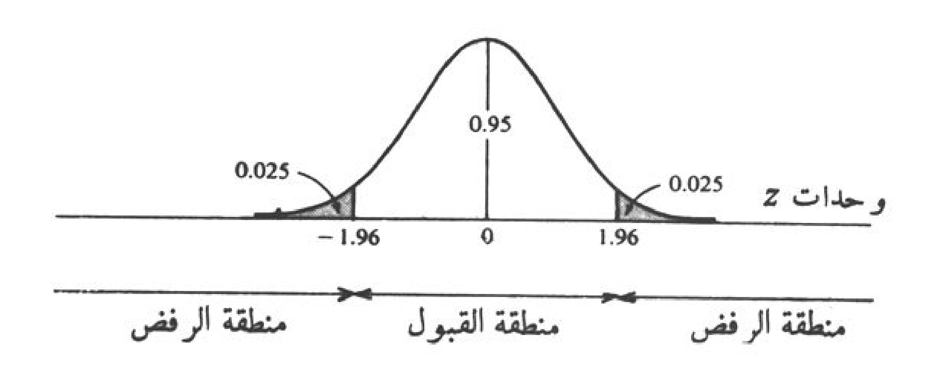
10.4 : تصنيف أنواع اختبار الفرضيات

تصنف اختبار الفرضيات بشكل عام الى اختبار في جانبين (طرفين) *two tailed test* واختبار في جانب واحد (طرف واحد) *one tailed test* . ومهما كان نوع الاختبار سوآءا كان ذو طرفين أو طرف واحد فأن فرضية العدم تكون واحدة مهما كانت الفرضية البديلة ، سوآءا كانت اقل او أكبر أو مجرد أنها تختلف .

وتنقسم منطقة رفض (الفرضية) ، وهي المساحة التي تعادل الخطأ (α) في الاختبار من طرفين الى قسمين واحدة الى اليمين والاخرى الى اليسار وكل منهما تساوي نصف مساحة الخطأ (α) أي α /2 . فاذا كانت مستوى الثقة 90% فهذا يعني أن المساحة تحت المنحنى لمنطقة القبول هي 0.90 وان مساحة منطقة الرفض ، التي تساوي الخطأ α هي 0.10 و أن مساحة كل من منطقتي الرفض على الطرفين هي 0.05 وتساوي α /2 ، وتصاغ الفرضية في حالة اختبار الفروض من طرفين كالاتي:

$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$

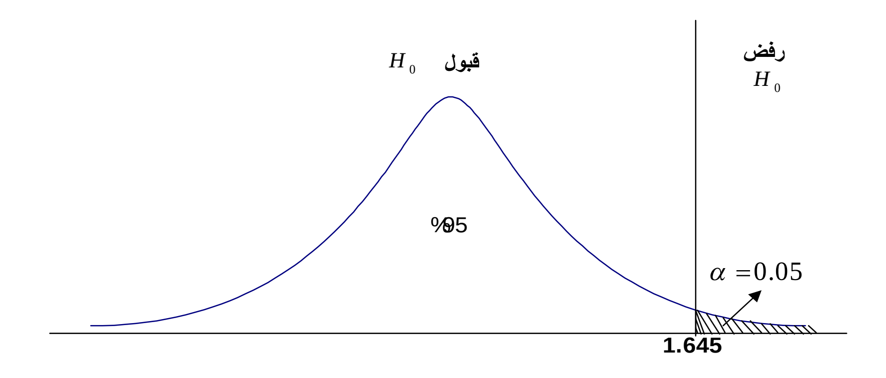


في الاختبار من طرف واحد ، فأن منطقة الرفض تقع في جانب واحد ، وتساوي مساحة الخطأ α وهي (0.10 مثلا) وتقع في الجانب الايمن أو الايسر . وتصاغ الفرضية في حالة اختبار الفروض من طرف واحد كالاتي :

اختبار الطرف الأيمن

$\mu_0 \quad =: \mu \ _0H$

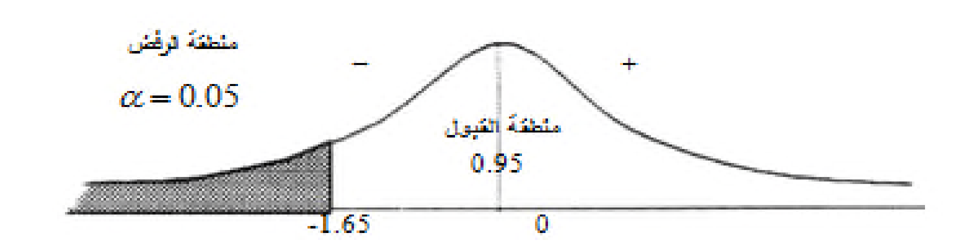
$\mu_0 > \quad : \mu \ _1H$



اختبار الطرف الايسر

$\mu_0 \quad =: \mu \ _0H$

$\mu_0 > \quad : \mu \ _1H$



9.5 اختبار الفروض للمتوسطات و النسب

10.5.1 : اختبار الفروض لمتوسط العينة [n ≥ 30]

تجرى اختبارات الفروض بشكل عام وفق الخطوات الاتية :

1. وضع فرضية العدم والفرضية البديلة .
2. تحديد مستوى الثقة (1-α) أو مستوى المعنوية (α) بهدف تحديد القيم الحرجة ، وهي القيم الجدولية لمنحنيات Z ، t ، .. F...ألخ .
3. حساب الاحصائي المناسب ، أي القيمة المحسوبة من البيانات أو المعيار الذي يتم في ضوئه قبول أو رفض فرضية العدم .
4. اتخاذ القرار حول متوسط العينة باستخدام قاعدة القرار من خلال مقارنة قيمة

الاحصائي المحسوب من البيانات بالقيم الحرجة الجدولية.

تختلف اختبارات الفروض حسب معرفة تباين المجتمع من عدمه ، طبيعة التوزيع و حجم العينة. ويستخدم توزيع Z في اختبار الفروض للمتوسطات عندما يكون التباين (الانحراف المعياري) للمجتمع معلوما والتوزيع طبيعيا ، مهما كان حجم العينة . لكن اذا كان التباين مجهول فيستخدم اختبار Z فقط عندما تكون العينة كبيرة [$n \geq 30$]

10.5.2 : خطوات اختبار الفروض لمتوسط العينة [$n \geq 30$]

- **وضع الفرضية ، وهي فرضية العدم والفرضية البديلة**

$H_0: \mu = \mu_0$ .... فرضية العدم

$H_1: \mu \neq \mu_0$ .... *الفرضية البديلة*

أو

$H_1: \mu > \mu_0$

$H_1: \mu > \mu_0$

- **تحديد مستوى المعنوية (α)** : تحدد مستوى المعنوية سلفا والتي تمثل الخطأ α وهي عادة 1% ، 5% أو 10% أو بمعنى أخر عند مستويات ثقة مقابلة (99% ، 95% و 90%) على التوالي. وفي ضوء هذا يتم تحديد قاعدة اتخاذ القرا ، اي تحديد القيمة الحرجة المعيارية (الجدولية) وهذه القيمة هي ($\pm Z_{\alpha/2}$) للاختبار من طرفين و($-Z_\alpha$) للاختبار من الطرف الايسر و ($+Z_\alpha$) للاختبار للطرف الايمن . ونستخدم جدول التوزيع الطبيعي المعيار **Z** للحصول على القيم الحرجة .
- نحسب القيمة المعيارية **Z** لمتوسط العينة أي الاحصائي المناسب أو القيمة المعيارية المحسوبة . وذلك باستخدام المعادلة الاتية :

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

- اتخاذ القرار حول متوسط العينة : إذا كانت القيمة المعيارية المحسوبة أقل من القيمة الجدولية ، فإنها بذلك تقع في منطقة القبول ونقبل فرضية العدم (**H0**) . وإذا كانت القيمة المعيارية المحسوبة أكبر من القيمة الجدولية ، فإنها بذلك تقع في منطقة الرفض ، ونرفض فرضية العدم ونقبل الفرضية البديلة (**H1**)

10.5.3 : اختبار الفروض لمتوسط العينة ($n < 30$)

عندما تكون العينة اقل من 30 ..($n<30$) مفردة والمجتمع التي اخذت منه العينة ذو توزيع طبيعي تقريبا والتباين (الانحراف المعياري) مجهول نستخدم اختبار t .

- **خطوات اختبار t**

- وضع الفرضية وهي فرضة العدم والفرضية البديلة

تحديد مستوى المعنوية (α) والذي تمثل الخطأ (α) مسبقا من بين مستويات المعنوية المختلفة (1% ، 5% او 10%) أو بمعنى اخر فأن مستويات الثقة المقابلة ($1-\alpha$) وهي (99% ، 95% و 90%) وفي ضوء هذا التحديد نحدد قاعدة اتخاذ القرار ، أي تحديد القيمة الحرجة المعيارية الجدولية وهذه القيمة هي ($\pm t_{\alpha/2}$) للاختبار من طرفين و ($- t_{\alpha}$ ، $+ t_{\alpha}$) للاختبار من الطرف الايسر والطرف الايمن على التوالي . ونستخدم جدول t للحصول على القيم الحرجة عند درجات حرية :($df = n-1$) .

- حساب القيمة المعيارية t لمتوسط العينة (القيمة المحسوبة) وذلك باستخدام المعادلة الأتية :

$$t=\frac{\overline{X}-\mu}{S/\sqrt{n}}$$

ونستخدم البيانات المأخوذة من العينة لحساب الانحراف المعياري لان قيمتها تكون مجهولة .

- أتخاذ القرار حول متوسط العينة باستخدام قاعدة القرار ، اذا وقعت القيمة المحسوبة في منطقة الرفض ، نرفض فرضية العدم وخلافا لذلك فأننا نقبلها .

10.5.4 :اختبار الفروض لنسبة واحدة للمجتمع [عينة كبيرة $n \geq 30$]

لا تختلف خطوات اختبار الفروض للنسب عن الخطوات المتبعة لاختبار الفروض للمتوسطات عندما تكون العينة كبيرة ، باستثناء اختلاف طريقة حساب الانحراف المعياري والخطأ المعياري للنسبة . ويرمز لنسبة المجتمع بالرمز (P) بينما يرمز لنسبة العينة بالرمز ($\overline{P}$) . وبافتراض أن توزيع المعاينة للنسبة للعينات الكبيرة هو توزيع طبيعي .

ويحسب الخطأ المعياري للنسبة وفقا للمعادلة التالية :

$$S_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

لا تختلف خطوات اختبار الفروض للنسب عن الخطوات المتبعة لاختبار الفروض لمتوسط العينة .

حساب القيمة المعيارية Z حسب المعادلة التالية

$$Z_{\bar{p}} = \bar{p} - p / S_{\bar{p}}$$

$$Z_{\bar{p}} = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

- ويلاحظ أننا نستخدم النسبة من المجتمع في حساب الخطأ المعياري للنسبة وليس النسبة من العينة . كما يلاحظ بأنه يتعين استخدام معامل التصحيح عند حساب الخطأ المعياري اذا تجاوز حجم العينة نسبة 5% من المجتمع .

10.6 : تطبيقات على اختبار الفروض للمتوسط والنسب

مثال(10.6.1): تقوم شركة للمواد الغذائية بإنتاج نوع من الشوكولاتة في عبوات 100 غرام . فاذا قامت وحدة الدراسات في الشركة اخذ عينة عشوائية من 36 عبوة لقياس الوزن للنظر فيما اذا كان وزن العبوات يختلف عن الوزن المعلن ، وبفرض أن الانحراف المعياري للمجتمع هو 10.5 غم . فاذا كان متوسط الوزن من العينة 110 غرام . استخدام مستوى ثقة 95% لاختبار الفرضية بأن متوسط العينة يختلف عن المتوسط المعلن .

الحل :

- وضع الفرضية

$$H_0: \mu = 100$$

$$H_1: \mu \neq 100$$

تم تحديد فرضية العدم والفرضية البديلة (حيث تستند الفرضية البديلة على ان المتوسط الفعلي لوزن العبوة قد يقل او يزيد عن المتوسط المعلن)

- **نحدد قاعدة أتخاذ القرار** ، اي القيمة الجدولية المعيارية Z عند مستوى ثقة 95% ، اي قيمة $Z_{\alpha/2}$ هي ± 1.96
- حساب القيمة المعيارية Z اي (القيمة المحسوبة Z)

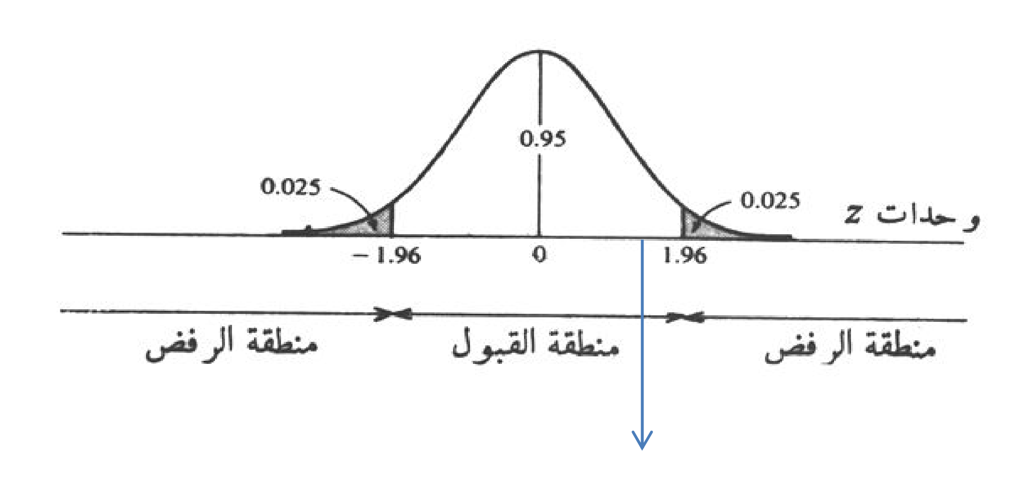
$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \quad .... \quad (\sigma_{\bar{x}} = \sigma/\sqrt{n})$$

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{(110 - 100)}{\sigma/\sqrt{n}}$$

$$= \frac{(53 - 50)}{10.5/\sqrt{36}}$$

$$= \frac{(3 \times 6)}{10.5} = 18 / 10.5 = 1.7$$



بما أن القيمة المحسوبة ($Z = 1.7$) أقل من القيمة الجدولية (1.96) فأننا نقبل فرضية العدم (H_0) . حيث أن البيانات المتوفرة لا تقدم دليلا على ان المتوسط المعلن غير صحيح . مما يشير أن الفرق بين متوسط العينة والمتوسط المعلن ، أنما يرجع للصدفة (وغير دال إحصائيا) .

مثال (10.6.2) إذا كان متوسط استهلاك اللحوم في احد البلدان 14 كغ عام 2000 ، فاذا أخذت عينة عشوائية من 49 فردا في عام 2010 ووجد من العينة ان متوسط استهلاك اللحوم للفرد من العينة 16 كغ والانحراف المعياري هو 6 كغ . أختبر الفرضية القائلة بأن متوسط الاستهلاك للحوم قد ارتفع عام 2010 مقارنة بعام 2000 وذلك عند مستوى ثقة 95% .

الحل :

- **وضع الفرضية**

$H_0 : \mu = 14$

$H_1 : \mu > 14$

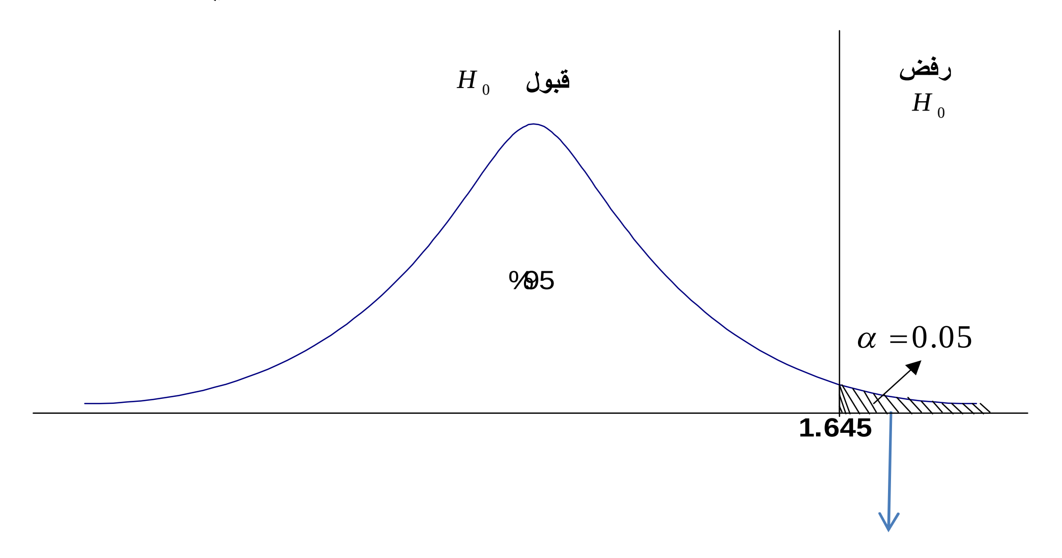
القيمة الجدولية المعيارية Z عند مستوى ثقة 95% ، اي قيمة ($Z_\alpha = 1.64$)

حساب القيمة المعيارية Z اي (القيمة المحسوبة Z)

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{(16 - 14)}{6/\sqrt{49}}$$

$$= \frac{2 X 7}{6} = 14/6 = 2.33$$



2.33

بما أن القيمة المحسوبة ($Z = 2.33$) اكبر من القيمة الجدولية ($Z = 1.645$) ، نرفض فرضية العدم ($H0$) ونقبل الفرضية البديلة ($H1$) . حيث أن البيانات المتوفرة تشير بأن استهلاك اللحوم للفرد عام 2010 تغير مقارنة بعام 2000 . والفرق بين متوسط العينة ومتوسط المجتمع جوهري .

مثال (10.6.3) يدعي وكيل سيارات في اعلان بأن السيارة التي يطرحها للبيع هي سيارة اقتصادية وتسير بالمتوسط 12 كم/ لكل لتر بنزين . وقد قامت جمعية حماية المستهلك بتحليلات احصائية باستخدام عينة عشوائية من 30 سيارة للنظر في صحة ادعاء وكيل السيارات ووجدت من العينة أن متوسط المسافة التي تقطعها السيارة هي 11 كم/ لكل لتر بنزين والانحراف المعيار ي 5 / لتر .

- أختبر الفرضية القائلة بأن متوسط المسافة التي تقطعها السيارة هو أقل مما يدعيه الوكيل باستخدام مستوى الثقة 90% .

الحل

- وضع الفرضية

$H_0 : \mu = 12$

$H_1 : \mu < 112$

- القيمة الجدولية المعيارية Z عند مستوى ثقة 90% ، اي قيمة (Z_α) التي تقع على يسارها 10% من المساحة وهذه القيمة هي ($Z = -1.28$)
- حساب القيمة المعيارية Z اي (القيمة المحسوبة Z)

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$Z = \frac{(12 - 11)}{5/\sqrt{30}}$$

$$= \frac{(-1) \times (\sqrt{30})}{5}$$

$$= -5.477 / 5 = -1.10$$

- بما أن القيمة المحسوبة (z = - 1.10) تقع في منطقة القبول ، نقبل فرضية العدم (H_0) ونرفض (H_1) ، حيث أن البيانات المتوفرة لا تقدم دليلا على أن المتوسط المعلن غير صحيح . مما يشير أن الفرق بين متوسط العينة والمتوسط المعلن انما يرجع الى الصدفة. وان المسافة التي تقطعها السيارة تتفق مع ادعاء الوكيل.

مثال (10.6.4) اختبار الذكاء (IQ) يتبع التوزيع الطبيعي بمتوسط 115 . فإذا تقدم لهذا الاختبار عينة تتكون من 25 طالبا بمدرسة معينة وكانت متوسط درجاتهم بهذا الاختبار يساوي 120 وتباين 100 . من هذه البيانات هل يمكن القول بأن متوسط درجات الطلبة بصفة عامة في هذه المدرسة يختلف عن المتوسط العام عند مستوى معنوية 5% .

الحل :

حيث أن حجم العينة اقل من 30 ($n < 30$) وتباين المجتمع مجهول ..فأننا نستخدم توزيع t

$H_0 : \mu = 115$

$H_1 : \mu \neq 115$

وحيث أن $\alpha = 0.05$ والاختبار من طرفين وعليه فأن ($t_{\alpha/2} = 0.025$) من جدول t عند درجات حرية ($df = n-1 = 25-1 = 24$) فالقيمة الجدولية الحرجة $t_{\alpha/2}$ هي 2.064

حساب القيمة المعيارية t اي (القيمة المحسوبة t)

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

$$t = \frac{(120 - 115)}{s/\sqrt{n}}$$

$$= \frac{(120 - 115)}{10/\sqrt{25}}$$

$$= \frac{5 \times 5}{10}$$

$$= 25/10 = 2.5$$

القيمة المحسوبة ($t = 2.5$) أكبر من القيمة الجدولية (2.064) وعليه نرفض فرض العدم (H_0) ونقبل الفرض البديل (H_1) والذي يدل على أن متوسط درجات الطلبة في هذه المدرسة يختلف عن المتوسط العام عند مستوى معنوية 5% .

مثال (10.6.5) أخذت عينة عشوائية من 50 عبوة من الارز من شحنة من 400 عبوة . ووجد ان عدد العبوات التي يقل وزنها عن الوزن المقرر هو 4 عبوات . فاذا كان المنتج يضمن أن لا تزيد نسبة العبوات الناقصة الوزن عن 3% .فهل يمكننا أن نستنتج بمستوى ثقة 95% أن ادعاء المنتج في غير محله .

الحل :

لا تختلف خطوات اختبار الفروض للنسب عن الخطوات المتبعة لاختبار الفروض لمتوسط العينة .

- وضع الفرضية :

$H_0 : P = 0.03$

$H_1 : P > 0.03$

- عند مستوى المعنوية (α =0.5) ، فأن القيمة الجدولية هي 1.64 وهو اختبار من الطرف الايمن والفرضية البديلة هي النسبة الفعلية أكبر مما يدعيه المنتج .

نحسب الخطأ المعياري للنسبة باستخدام معامل التصحيح لان حجم العينة أكبر من 5% من حجم المجتمع

$n/N = 50/400 = 0.125 > 0.05$

$$S_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$S_{\bar{p}} = \sqrt{\frac{(0.03)(0.97)}{50}} \cdot \sqrt{\frac{400-50}{400-1}}$$

$$S_{\bar{p}} = 0.024 \text{ x } 0.936 = 0.0225$$

- حساب القيمة المعيارية Z حسب المعادلة التالية

$$Z_{\bar{p}} = (\bar{p} - p)/ S_{\bar{p}}$$

$$Z_{\bar{p}} = (\bar{p} - p) / S_{\bar{p}} \qquad (\bar{p} = 50/400 = 0.08\ 0)$$

$$= (0.08 - 0.03) / 0.0225 = 2.2$$

نجد أن القيمة المحسوبة تقع في منطقة الرفض ولذلك نرفض فرضية العدم (H_O) ونقبل الفرضية البديلة (H_1) حيث أن قيمة z المحسوب (2.2) أكبر من القيمة الجدولية (1.64) ونستنتج من ذلك بأن البيانات المتوفرة تقدم دليلا على ان ادعاء المنتج غير صحيح لان الفرق بين النسبة المشاهدة والنسبة التي يدعيها المنتج كبير وجوهري من ان يكون ناتجا عن الصدفة وحدها . أي أن الفرق جوهري إحصائيا .

مثال (10.6.6) ترغب شركة أن تعرف بدرجة ثقة 95% ما اذا كان يمكن الادعاء بأن صناديق الصابون المسحوق الذي تبيعه يحتوي على أكثر من 400 جرام من الصابون . وتعرف الشركة من الخبرة الماضية أن أوزان الصابون بالصناديق يتبع التوزيع الطبيعي . وقد اخذت الشركة عينة عشوائية حجمها 25 ووجدت أن الوسط الحسابي 420 جرام والانحراف المعياري 75 .

الحل :

$H_0 : \mu = 400$

$H_1 : \mu > 400$

حيث ان التوزيع طبيعي ($n < 30$) وكذلك σ غير معلومة ، نستخدم توزيع t ($df = n-1 = 25-1 = 24$) تحديد المنطقة الحرجة للاختبار وبمستوى معنوية 0.05 .

حساب القيمة المعيارية t اي (القيمة المحسوبة t)

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

$$t = \frac{(420 - 400)}{75/\sqrt{25}}$$

$$= \frac{(20) \times 5}{75}$$

$$= \frac{20}{15}$$

$$= 1.33$$

قيمة t الجدولية عند درجات الحرية 24 = 1.711 و نقارنها مع قيمة t المحسوبة = 1.33

بما أن قيمة t المحسوبة (1.33) اقل من قيمة (t) الجدولية (1.711) نقبل H_0 ونرفض H_1 وهذا يعني بأن ادعاء الشركة غير صحيح .

❖ أختبار الفروض للفرق بين متوسطين ($n_1 , n_2 \geq 30$)

➢ خطوات الاختبار

- وضع الفروض وهي فرضية العدم والفرضية البديلة

$$Ho: \mu_1 = \mu_2$$

$$HI: \mu_1 \neq \mu_2$$

- تحديد مستوى المعنوية (α) أي تحديد قيمة الخطأ α (عادة 1% ، 5% ، 10%) ومن ثم تحديد القيمة الحرجة الجدولية
- حساب القيمة المعيارية (Z) لمتوسط العينة (القيمة المحسوبة) والاحصائي المستخدم في هذا الاختبار هو اختبار Z

$$Z_{\bar{X}1-\bar{X}2} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

حيث أن :

$\sigma_{\bar{X}_1-\bar{X}_2}$ = الخطأ المعياري للفرق بين متوسطين ويحسب كالاتي:

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

وتستخدم قيم الانحراف المعياري S_1^2 ، S_2^2 المحسوبة من العينة التي تساوي أو تزيد عن 30 مفردة ، وهي تقديرات جيدة للانحراف المعياري للمجتمعين ، ولكن في حالات نادرة تكون فيها الانحرافات المعيارية لكلا المجتمعين معروفة فأنه يتعين استخدامها .

أتخاذ القرار حول ما اذا كان الفرق بين المتوسطين جوهريا أم لا من خلال مقارنة قيمة الاحصائي المحسوب من العينات مع القيمة الجدولية

مثال (10،6،7) : يرغب مدير أن يحدد عند مستوى معنوية 5% ما أذا كان الاجر بالساعة($) للعمال نصف المهرة متساويان في مدينتين ، لعمل ذلك يأخذ عينة عشوائية من الاجر بالساعة من كل من المدينتين وكانت النتائج كالاتي :

$\bar{X}_1 = 6.00$ ، $\bar{X}_2 = 5.40$

$S_1 = 2.00$, $S_2 = 1.80$

$n_1 = 40$, $n_2 = 54$

الحل :

$$Ho: \mu_1 = \mu_2$$

$$HI: \mu_1 \neq \mu_2$$

عند مستوى معنوية 5% فأن القيمة الجدولية $Z_{\alpha/2} = \pm 1.96$

نحسب القيمة المعيارية Z (القيمة المحسوبة)

نحسب أولا الخطأ المعياري $\sigma_{\bar{X}_1-\bar{X}_2}$

$$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}=\sqrt{\frac{2.00^2}{40}+\frac{1.80^2}{54}}=\sqrt{0.1+0.06}=\sqrt{0.16}=0.4$$

$$z==\frac{(\bar{X}_1-\bar{X}_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}=\frac{0.6}{0.4}=1.5$$

وحيث أن قيمة z المحسوبة(1.5) اقل من القيمة الجدولية(1.96) و تقع داخل منطقة القبول ، فأننا نقبل H_0، أي $\mu_1 = \mu_2$ ، عند مستوى معنوية 5%.

❖ اختبار الفروض للفرق بين متوسطين لعينات صغيرة (n1, n2 < 30)

الفرضيات :

I. العينات صغيرة

II. توزيع المجتمعين طبيعي

III. الانحراف المعياري للمجتمعين غير معلوم لكنه متساو

نستخدم توزيع t بدرجات حرية ($df = n_1 + n_2 - 2$)

$$t==\frac{(\bar{X}_1-\bar{X}_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}$$

الخطأ المعياري للفرق بين متوسطين يحسب كالاتي :

$$S_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{s^2}{n_1}+\frac{s^2}{n_2}}$$

التباين المستخدم في المعادلة هو **التباين التجميعي** (التراكمي) ويحتسب من كل من العينتين حسب المعادلة الاتية:

$$S^2 = \frac{S_1^2\,(n_1-1)+S_2^2\,(n_2-1)}{n_1+n_2-2}$$

مثال (10.6.8) تقوم أحدى الشركات بإنتاج منتج جديد من مادة تنظيف . ويرغب مدير التسويق في الشركة في مقارنة متوسطات مبيعات الشركة الشهرية في منطقتين 1 ، 2 باستخدام عينة عشوائية من المحلات التجارية في المنطقتين ، أختبر الفرضية القائلة بأن هناك فرقا بين متوسطات المبيعات في المنطقتين عند مستوى معنوية 5% وذلك من نتائج العينة الاتية :

$\bar{X}_1 = 40 \quad , \quad \bar{X}_2 = 38$

$S_1 = 15.46\,, \ S_2 = 4.34 \qquad (S_1^2 = 239.011, \ \ S_2^2 = 18.84)$

$n_1 = 10 \quad , n_2 = 15$

الحل :

$\mathbf{H_0} : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

عند مستوى ثقة 95% (مستوى المعنوية 5% وهو الخطأ α) فأن القيمة الجدولية لبقية الحرجة t عند درجات حرية 23 هي 1.71

$(df = n_1 + n_2 - 2 = 10 + 15 - 2 = 23)$.

نوجد التباين التجميعي (التراكمي)

$$S^2 = \frac{S_1^2\,(n_1 - 1) + S_2^2\,(n_2 - 1)}{n_1 + n_2 - 2}$$

$$S^2 = \frac{239.01(10-1) + 18.84(15 - 1)}{10+15-2} = 103.8$$

$S^2 = 103.8$

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{103.8}{10} + \frac{103.8}{15}} = \sqrt{10.38 + 6.92} = 4.16$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{(40-38)}{4.16} = 0.48$$

قيمة t المحسوبة اقل من قيمتها الجدولية. وتقع القيمة المحسوبة للإحصائي (t) في منطقة القبول ، وبذلك نقبل فرضية العدم H0 . و نستنتج من ذلك أن البيانات المتوفرة لا تقدم دليلا كافيا على ان هناك فرق جوهري بين متوسطي عدد العبوات المباعة للمحلات التجارية في المنطقتين.

❖ **اختبار الفروض للفرق بين نسبتين**

كما أن هناك توزيعات للمعاينة بين المتوسطات أو ألنسب والفروق بين المتوسطات ، هناك أيضا توزيعات للمعاينة للفروق بين النسب تسمح بتقدير الخطأ المعياري للفرق بين نسبتين ومن ثم تقدير الاحتمالات لحصول فروق معينة .

معادلة الخطأ المعياري للفرق بين نسبتين هي:

$$\sigma_{\bar{p}_1-\bar{p}_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}$$

ويحسب الخطأ المعياري للفرق بين نسبتين باستخدام النسبة العام ($\bar{P}$) وفقا للمعادلة التالية :

وحتى يمكن تقدير ما اذا كانت الفروق بين النسب راجعة للصدفة أو هي فروق جوهرية فأننا نحول قيمة المتغير العشوائي (الفرق بين نسبتين) الى قيمة معيارية ، والقيمة المعيارية الذي يمثل قيمة الاحصائي المستخدم في اختبار الفروض هي :

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}$$

$$Z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sigma_{\bar{p}_1 - \bar{p}_2}}$$

مثال (10.6.9) : ترغب شركة أن تحدد عند مستوى معنوية 1% ما إذا كانت نسبة المقبول من المكونات الإلكترونية لمورد أجنبي ، p_1 تزيد عنها لمورد محلي، p_2. وقد أخذت الشركة عينة عشوائية من شحنة كل مورد ووجدت أن $p_1 = 0.9$ و $p_2 = 0.7$ من عينات من حجم $n_1 = 100$ و $n_2 = 80$.

الحل

الفروض :

$H_0 : p_1 = p_2$

$H_1 : p_1 > p_2$

قيمة z_α عند مستوى المعنوية 1% من جدول التوزيع الطبيعي المعياري هي (2.58)

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{(100)(0.9) + (80)(0.7)}{180} = \frac{146}{180} = 0.8$$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}} = \sqrt{\frac{(0.8)(0.2)}{100} + \frac{(0.8)(0.2)}{80}} = \sqrt{0.0016 + 0.002} = \sqrt{0.0036} = 0.06$$

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sigma_{\bar{p}_1 - \bar{p}_2}} = \frac{0.9 - 0.7}{0.06} = 0.2 / 0.06 = 3.33$$

بما أن القيمة المحسوبة للإحصائي z_α (3.33) أكبر من القيمة الجدولية (2.58) نرفض H_0 ونقبل الفرض H_1 أي أن $p_1 > p_2$ عند مستوى معنوي 1% .

10.7 : اختبار كاى - تربيع لجودة التوفيق والاستقلال

توزيع كاي تربيع (x^2) هو أحد التوزيعات الاحتمالية المتصلة. ومن بين استخدامات هذا الاختبار ، اختبارات الفروض المتعلقة بجودة التوفيق والاستقلالية (المتعلق بتحديد استقلالية خاصيتين أو أكثر) لمجتمع ما . ويستخدم هذا الاختبار في اختبارات الفروض للبيانات الموزعة في فئات . ولا يستند هذا الاختبار على فرضية التوزيع الطبيعي لمجتمع الدراسة شأن اختبارات (Z ، T ، F) والتي تسمى بالاختبارات المعلمية (PARAMETRIC TEST) . ولذلك فأن اختبار (x^2) يعرف بأنه من الاختبارات اللامعلمية (NON PARAMETRIC TET) . وكما هو الحال في التوزيع الطبيعي فأن الاحتمالات للمتغير العشوائي الذي له توزيع (x^2) تساوي المساحات تحت منحنيات (x^2) . ويمكن ايجادها باستخدام جدول خاص بهذا التوزيع . ويسمح جدول (x^2) بإيجاد قيم كاي تربيع عند درجات حرية معينة .

منحنى كاي تربيع يتوقف التوائه على درجات الحرية. أي أن هناك علاقة طردية بين قيم (x^2) ودرجات الحرية .

شكل 10.1 : منحنى كاي تربيع

خصائص منحنى كاي تربيع

1. المساحة الكلية تحت المنحنى = 1
2. يبدأ المنحنى من نقطة الصفر ويمتد الى مالا نهاية الى اليمين مقتربا من المحور الافقي من دون ان يلامسه .
3. المنحنى غير متماثل ويلتوي الى اليمين فترتفع بسرعة الى القمة ثم يخدر ببطئ اتجاه اليمين وكلما زاد عدد درجات الحرية فان المنحنيات تأخذ شكل المنحنى الطبيعي بشكل متزايد.

❖ **حساب (x^2) لجودة التوفيق (GOODNESS OF FIT)**

يستخدم توزيع كاي – تربيع (x^2) لاختبار :

(1) إذا كانت التكرارات المشاهدة تختلف (معنوياً) عن التكرارات المتوقعة عندما يكون عدد النواتج الممكنة أكثر من أثنين،

(2) إذا كان التوزيع الذي آخذت منه العينة ذا الحدين، أو الطبيعي، أو أي توزيع أخر، (3) إذا كان متغيران مستقلين أم لا.

وإحصائية x^2 المحسوبة من بيانات العينة معطاة بالصيغة :

$$x^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

حيث f_0 = التكرارات المشاهدة

f_e = التكرارات المتوقعة

فإذا كانت x^2 المحسوبة أكبر من القيمة الجدولية x^2 عند مستوى المعنوية ودرجات الحرية المحدودة ، يرفض الفرض العدمي H_0 لصالح الفرض البديل H_1 .

درجات الحرية لاختبار جودة التوفيق معطاة بالصيغة :

$$df = c - m - 1$$

حيث أن : c = عدد الفئات

m = عدد معالم المجتمع التي يجري تقديرها من إحصائيات العينة.

درجات الحرية لاختبارات الاستقلال أو اختبارات جداول الاقتران معطاة بالصيغة :

$$df = (r - 1)(c - 1)$$

حيث r = عدد الصفوف في جدول الاقتران

n = عدد الأعمدة.

ويكون التكرار المتوقع فى كل خلية من جدول الاقتران

$$f_e = \frac{\sum r \sum c}{n}$$

حيث n = حجم العينة الإجمالي.

مثال (10.7.1) : وجد محل تجاري من خبرته الماضية أن 30% من التليفزيونات المباعة من الحجم الصغير ، 40% من الحجم المتوسط، 30% من الحجم الكبير. لتحديد جم المخزون الواجب الاحتفاظ به من كل نوع، أخذ المدير عينة عشوائية من 100 من المبيعات الحديثة للتليفزيون فوجد أن منها 20 من النوع الصغير، 40 من النوع المتوسط، 40 من النوع الكبير. باستخدام مستوى معنوية 5%، أختبر الفرضية القائلة أن نمط المبيعات في الماضي لازال سائداً .

الحل :

الفرضيات :

نمط المبيعات في الماضي لازال سائدا : H_0

يوجد اختلاف في نمط في نمط المبيعات مقارنة بالماضي : H_1

جدول (10.1) المشتريات المشاهدة والمتوقعة لأجهزة التليفزيون حسب جم الشاشة

الإجمالي	جم الشاشات			
	صغير	متوسط	كبير	
100	40	40	20	النمط المشاهد f_0
100	30	40	30	f_Eالنمط في الماضي

$$df = c - m - 1 = 3 - 0 - 1 = 2$$

قيمة x^2 الجدولية عند مستوى معنوية 5% بدرجة حرية 2 هي(5.99)

$$x^2 = \sum\frac{(f_0 - f_e)^2}{f_e} = \frac{(20-30)^2}{30} + \frac{(40-40)^2}{40} + \frac{(40-30)^2}{30} = \frac{-10^2}{30} + \frac{0^2}{40} + \frac{10^2}{30} = \frac{100}{30} + \frac{100}{30} = 6.66$$

وحيث أن القيمة المحسوبة $x^2 = 5.83$ أصغر من القيمة الجدولية $x^2 = 5.99$ بمستوى معنوية $\alpha = 0.05$ ودرجات حرية 2 فإننا نرفض H_o ،ونقبل H_1 البديلة . أي يوجد اختلاف في نمط المبيعات في التلفزيونات حاليا مقارنة بما كان سائدا في الماضي .

مثال (10.7.2): جمع تاجر سيارات البيانات الموضحة في جدول (10.2) عن عدد السيارات الأجنبية والمحلية التي يشتريهما عملاء أعمارهم تحت سن 30 سنة، والتي يشتريها عملاء أعمارهم

سن 30 سنة فأكثر. أختبر ما إذا كان نوع السيارة المشتراه (أجنبية أو محلية) مستقبلاً عن سن المشتري 1 عند معنوية %.

جدول (10.2) جدول الاقتران لمشتري السيارات

الإجمالي	نوع السيارة		السن
	أجنبية	محلية	
70	40	21	تحت 30 ،
100	80	29	30 ، فأكثر
170	120	50	إجمالي

الحل

الفرضيات

نوع السيارة المشتراه (أجنبية أو محلية) مستقل عن سن المشتري : H_o

نوع السيارة المشتراه (أجنبية أو محلية) ليس مستقلا عن سن المشتري : H_1

ننشئ جدول التكرارات المتوقعة (جدول 10.3) القيمة في الخلية الأولى صف 1 وعمود 1

$$f_e = \frac{\sum r \sum c}{n} = \frac{(70)(50)}{170} \cong 21$$

ويمكن الحصول على قيم بقية الخلايا بنفس الطريقة

جدول (10.3) جدول التكرارات المتوقعة المناظرة للتكرارات

الإجمالي	نوع السيارة		السن
	أجنبية	محلية	
70	49	21	تحت 30 ،
100	71	29	30 ، فأكثر

إجمالي	50	120	170

$$df = (r-1)(c-1) = (2-1)\ (2-1) = 1$$

$$x^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(30-21)^2}{21} + \frac{(40-49)^2}{49} + \frac{(20-29)^2}{29} + \frac{(80-71)^2}{71} = 9.44$$

وحيث أن قيمة x^2 المحسوبة(9.44) أكبر من القيمة x^2 الجدولية (6.66) وذلك عند α = 0.01 و 1 = df ، و لهذا نرفض H0 القائل بأن السن ليس عاملاً في تحديد نوع السيارة المشتراه ونقبل H_1 والذي يعني أن نوع السيارة المشتراه (أجنبية أو محلية) ليس مستقلا عن سن المشتري (وتنتهي إلى أن الأصغر سنناً يميلون فيما يبدو إلى شراء السيارات الأجنبية).

10.8 : اختبار الفروض لأكثر من متوسطين (تحليل التباين)

Analysis of variance (ANOVA)

يستخدم تحليل التباين لاختبار فرض أن متوسطات أكثر من مجتمعين متساوية أو مختلفة عندما تكون المجتمعات موزعة توزيعاً طبيعياً مع تساوي التباين. وتعود تسمية تحليل التباين الى أن إجراءات المقارنة بين المتوسطات تتضمن تحليلا للتباين بين بيانات العينة . ويستهدف التحليل تبيان فيما اذا كانت متوسطات المجتمعات متساوية تقريبا وأن أي اختلافات بينهما أنما تعود للصدفة ويمكن توقعها من عينات عشوائية من مجتمع توزيعه طبيعي ، أم ان أن التباين ناتج من اختلاف المتوسطات (غير متساوية) والفروق بينهما جوهرية .

- **خطوات أجراء تحليل التباين**

1. نقدر تباين المجتمع من التباين بين متوسطات العينات (MSA)
2. نقدر تباين المجتمع من التباين داخل العينات (MSE)
3. نحسب النسبة F (MSA / MSE)

$$F = \frac{\text{التباين بين متوسط العينات}}{\text{التباين داخل العينات}}$$

إذا كانت F المحسوبة أكبر من قيمة F الجدولية عند مستوى المعنوية ودرجات الحرية المعينة فإن الفرض العدمي H_0 عن تساوي متوسطات المجتمعات، يرفض لصالح الفرض البديل، H_1.

جدول (10.4) جدول تحليل التباين

النسبة F	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر التغير
$\frac{MSA}{MSE}$	$MSA=\frac{SSA}{c-1}$	$c-1$	$SSA=r\sum(\bar{X}_J-\bar{\bar{X}})^2$	بين الأوساط (يفسره العامل A)
--	$MSE=\frac{SSE}{(r-1)c}$	$(r-1)c$	$SSE=\sum\sum(\bar{X}_{iJ}-\bar{\bar{X}}_j)^2$	داخل العينات (الخطأ أو غير المفسر)
--	--	$rc-1$	$SST=\sum\sum(X_{iJ}-\bar{\bar{X}})^2=SSA+SSE$	الإجمالي

حيث أن :

$\bar{X}_J=$ متوسط العينة J المكونة من r مشاهدة = $(\Sigma_i X_{iJ}) / r$

$\bar{X}=$ المتوسط الكبير لكل العينات $(\Sigma_i \Sigma_J X_{iJ})/ rc$

SSA= مجموع المربعات التي يفسرها العامل $A = r\Sigma (X_J - \bar{X})^2$

SSE= مجموع مربعات الخطأ والتي لا يفسرها العامل $A = \Sigma\Sigma (X_{iJ} - \bar{X}_J)$

SST= مجموع المربعات الإجمالي = $SSA + SSE = \Sigma\Sigma (X_{iJ} - \bar{X})^2$

ويعطي قيم F عندما $\alpha - 0.05$ (الرقم الأعلى) وعندما $\alpha = 0.01$ (الرقم الأسفل) لكل زوج من درجات الحرية :

درجات حرية البسط = $c - 1$

حيث c = عدد العينات

درجات حرية المقام = $(r - 1) c$

حيث r = عدد المشاهدات في كل عينة .

مثال(10.8): تبيع شركة نفس الصابون في ثلاثة أغلفة مختلفة وبنفس السعر. يبين الجدول أدناه مبيعات 5 شهور. المبيعات موزعة توزيعاً طبيعياً ولها تباين متساو. أختبر ما إذا كان متوسط المبيعات لكل غلاف متساوياً أم لا عند مستوى معنوية 5%

غلاف (1)	غلاف (2)	غلاف (3)
87	78	90
83	81	91
79	79	84
81	82	82
80	80	88
ــــ	ــــ	ــــ
410	400	435

الحل

$H_0: \mu 1 = \mu_2 = \mu_3$

$H_1 \neq \mu_1 , \mu_2 , \mu_3$

$\overline{X}_1 = \frac{410}{5} = 82$, $\overline{X}_2 = \frac{400}{5} = 80$., $\overline{X}_3 = \frac{435}{5} = 87$

$\overline{X} = \frac{410 + 400 + 435}{(5)(3)} = 83$

$SSA = 5[(82-83)^2 + (80-83)^2 + (87-83)^2] = 130$

$SSE = (87-82)^2 + (83-82)^2 + (79-82)^2 + (81-82)^2 + (80-82)^2 + (78-80)^2 + (81-80)^2 + (79-80)^2 + (82-80)^2 + (80-80)^2 + (90-87)^2 + (91-87)^2 + (84-87)^2 + (82-87)^2 + (88-87)^2 = 110$

$$SST = (87-83)^2 + (83-83)^2 + . . + (88-83)^2 = SSA + SSE = 240$$

وتستخدم البيانات السابقة لتكوين جدول (10.5) لتحليل التباين ANOVA

جدول (10.5) جدول ANOVA لأغلفة الصابون

التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة
MSA/MSE = 65/9.17 =7.09	MSA = 130/2 = 65	$c - 1 = 2$	SSA = 130	تفسره الأغلفة (بين الأعمدة)

الخطأ أو غير المفسر (داخل الأعمدة	SSE = 110	$(r-1)c = 12$	MSE = 110/12 = 9.17	
الإجمالي	SST = 240	$Rc - 1 = 14$	--	

وحيث أن القيمة المحسوبة F = 7.09 (من جدول 10.5) أكبر من القيمة الجدولية F = 3.88 عند $\alpha = 0.05$ ودرجات حرية 2 و 12 ، فإننا نرفض H0 ، أي الفرض القائل بأن متوسط المبيعات للأغلفة المختلفة متساوية ، وتقبل H1 ، بأنها تختلف . ويشار إلى الإجراء السابق بأنه تحليل التباين في اتجاه واحد أو لعامل واحد.

تمارين الفصل العاشر

1. ترغب شركة ان تعر بدرجة ثقة 95% ما اذا كان يمكنها الادعاء بأن صناديق الصابون المسحوق الذي تبيعه يحتوي على اكثر من 500 جرام من الصابون . وتعرف الشركة من الخبرة الماضية لأن اوزان الصابون بالصناديق تتبع التوزيع الطبيعي ، وقد أخذت الشركة عينة عشوائية حجمها 25 مردة ووجدت ان الوسط الحسابي يساوي 520 جرام و الانحراف المعياري 75 . كي يمكن صياغة الفرضية واثبات ادعاء الشركة .
2. يرغب منتج كابلات من الصلب اختبار ما اذا كانت الكابلات التي ينتجها لديها قوة مقاومة للكسر قدرها 5000 رطل . قوة المقاومة للكسر اقل من 5000 رطل لن تكون ملائمة ، وقوة المقاومة للكسر أكبر من 5000 ترفع التكاليف بدون مبرر . يأخذ المنتج عينة عشوائية من 64 قطعة ويجد ان متوسط قوة المقاومة للكسر هو 5100 والانحراف المعياري هو 480 . هل يجب ان يقبل المنتج الفرض بأن كابلات الصلب لها قوة مقاومة للكسر 5000 عند مستوى معنوية 5% .
3. يعرف مركز تجنيد بالجيش بالخبرة بالماضية أن وزن المجند يتبع التوزيع الطبيعي بوسط حسابي يساوي 80 كغ وانحراف معياري 10 . ويرغب المركز أن يختبر عند مستوى معنوية 1% ما اذا كان متوسط وزن مجندي هذا العام أكبر من 80 كغ ، ولعمل هذا فقد أخذ عينة عشوائية من 25 مجندا حيث وجد ان متوسط الوزن في العينة 85 كغ .كيف يمكن اجراء هذا الاختبار .
4. يريد مستشفى ان يختبر ان 90% من جرعات عقار يشتريه يحتوي على 100 ملغ من العقار . لعمل هذا يأخذ المستشفى عينة من 100 مفردة (جرعة) ويجد ان 85 منها

فقط يحتوي على الكمية المناسبة . كيف للمستشفى ان يختبر هذا عند مستوى معنوية 5% و 10% .

5. نوعين من المصابيح الكهربائية تم اختبارها للتعرف على عمرها التشغيلي وكانت النتائج كالتالي :

$\bar{x}_1 = 1234$ (ساعة احتراق) .... $S_1 = 36$... $n_1 = 8$

$\bar{x}_2 = 1136$ ... $S_2 = 40$ ... $n_2 = 7$

هل الفرق بين المتوسطين جوهري عند مستوى معنوية 5% .

6. عينتان حجماهما 6 و 5 مفردات على التوالي اعطتا البيانات التالية :

$\bar{x}_1 = 40$.... $S_1 = 8$

$\bar{x}_2 = 50$ ... $S_2 = 10$

هل الفرق جوهري بين متوسطي العينتين عند مستوى معنوية 5%

7. في عينة مكونة من 1000 شخص من مدينة معينة ، وجد ان 450 منهم يدخنون ، و من عينة مكونة من 800 شخص من مدينة اخرى وجد ان 400 منهم يدخنون . هل البيانات المعطاة تؤكد بأن المدينتين تختلفان جوهريا فيما يتعلق عادة التدخين وذلك عند مستوى معنوية 5% .

8. يعطي جدول الاقتران أدناه ، عدد الاجزاء الالكترونية المقبولة وغير المقبولة المنتجة خلال ساعات الصباح المختلفة في عينة عشوائية من أنتاج المصنع . هل يجب قبول أو رفض الفرض عند مستوى معنوية 5% بأن أنتاج الوحدات المقبولة مستقل عن ساعة الصباح التي انتج خلالها .

	9- 8 صباحا	10- 9 صباحا	11- 10 صباحا	12 – 11 صباحا	الاجمالي
مقبولة	60	75	80	65	280
غير مقبولة	30	25	30	35	120
أجمالي	90	100	110	100	400

9. أعطت عينة عشوائية من 37 عاملا فوق سن 65 في مدينة ما النتائج الواردة في جدول الاقتران أدناه ، باستخدام مستوى المعنوية 10% أختبر الفرض بأن عدد الاناث

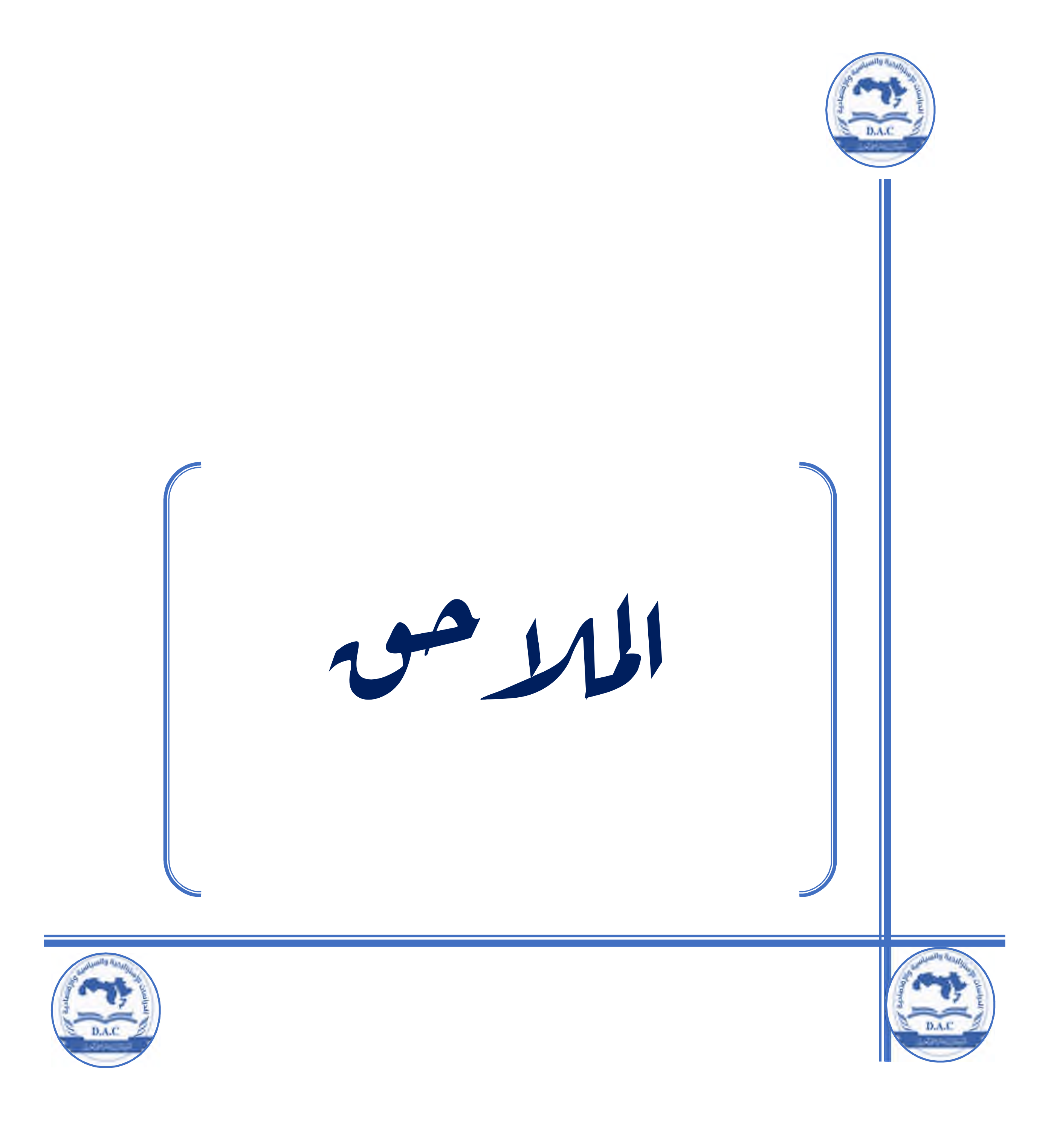
والذكور من العاملين في مجموعات السن 66-70 ، 71 + ، في المدينة مستقل عن الجنس .

أجمالي	ذكور	أناث	فئة العمر
26	9	17	70-66
11	8	3	71+
37	17	20	اجمالي

10.الجدول ادناه يبين عدد الاميال في الجالون لا ربعة انواع من البنزين لمدة 5 ايام . أفترض ان عدد الاميال للجالون لكل نوع موزع توزيعا طبيعيا مع تساوي التباين . هل يجب قبول او رفض الفرض بأن المتوسطات للمجتمع متساوية عند مستوى معنوية 5%.

عدد الاميال للجالون لا ربعة أنواع من البنزين لمدة 5 ايام

النوع الرابع	النوع الثالث	النوع الثاني	النوع الاول
17	16	12	12
15	14	14	11
17	15	13	12
16	13	15	13
18	14	14	11


الملاحق

ملحق (1) جدول التوزيع الطبيعي المعياري (Z)

Areas Under the One-Tailed Standard Normal Curve

This table provides the area between the mean and some Z score.
For example, when Z score = 1.45 the area = 0.4265.

σ = 1
0.4265
z
μ=0
1.45

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

ملحق (2) : جدول توزيع t

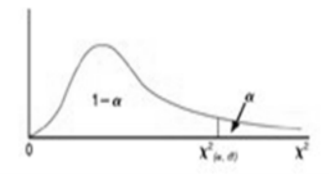
STUDENT t DISTRIBUTION

The first column lists the number of degrees of freedom (k). The headings of the other columns give probabilities (P) for t to exceed the entry value. Use symmetry for negative t values.

Two-tailed Conf. Level :	.80	.90	.95	.98	.99
df \ P	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

Source: Donald J. Koosis, Business Statistics *(New York: John Wiley & Sons, 1972). Reprinted by permission.*

ملحق (3) جدول مربع كاي



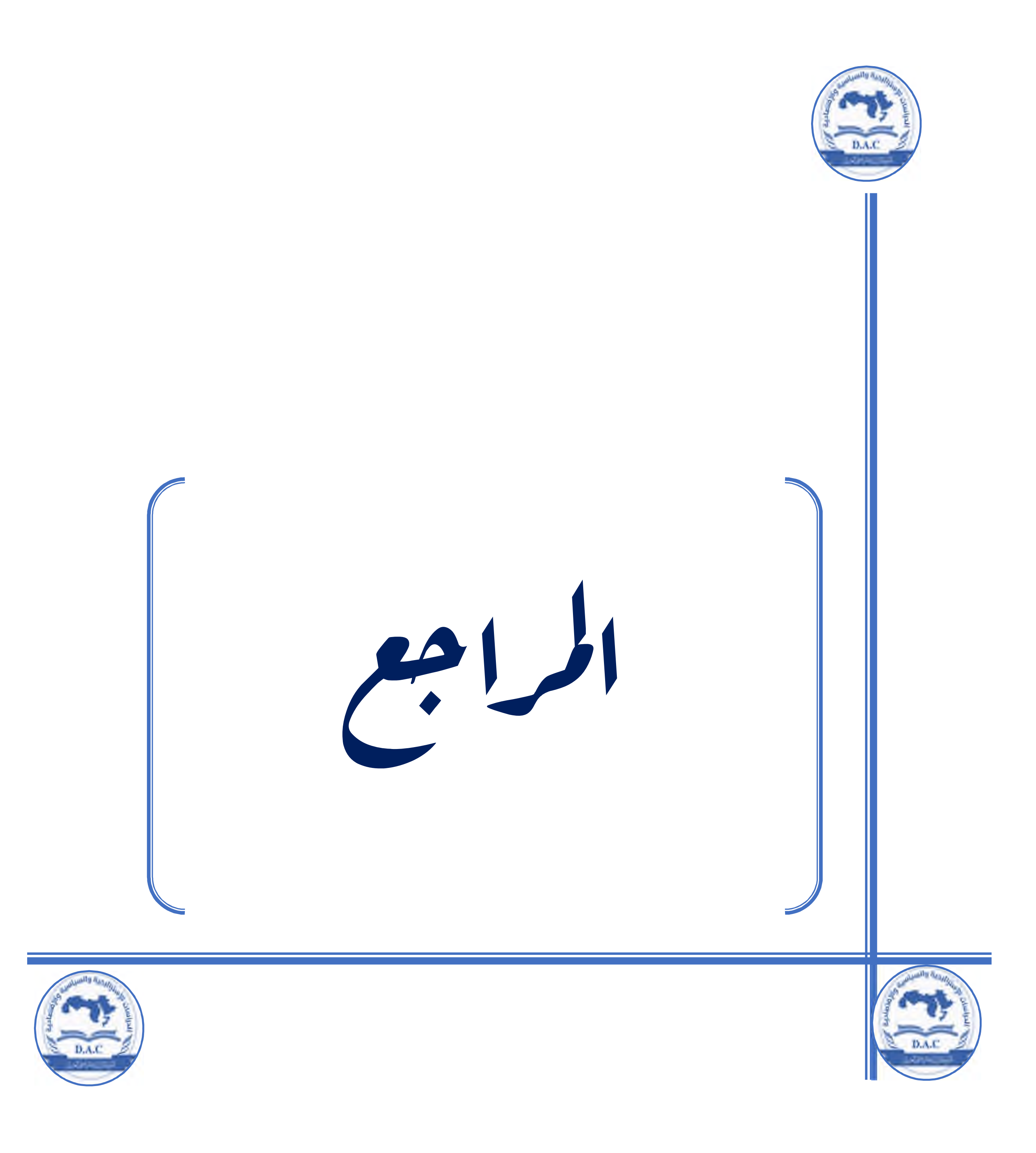
Degrees of Freedom	Upper Tail Areas (α)											
	.995	.99	.975	.95	.90	.75	.25	.10	.05	.025	.01	.005
1			0.001	0.004	0.016	0.102	1.323	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	0.575	2.773	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	1.213	4.108	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	1.923	5.385	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	2.675	6.626	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	3.455	7.841	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	4.255	9.037	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	5.071	10.219	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	5.899	11.389	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	6.737	12.549	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	7.584	13.701	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	8.438	14.845	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	9.299	15.984	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	10.165	17.117	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	11.037	18.245	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	11.912	19.369	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	12.792	20.489	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	13.675	21.605	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	14.562	22.718	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	15.452	23.828	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	16.344	24.935	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	17.240	26.039	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	18.137	27.141	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	19.037	28.241	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	19.939	29.339	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	20.843	30.435	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	21.749	31.528	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	22.657	32.620	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	23.567	33.711	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	24.478	34.800	40.256	43.773	46.979	50.892	53.672

الأستاذ الدكتور / علي أحمد السقاف

ملحق (4) جدول توز

F - Distribution ($\alpha = 0.05$ in the Right Tail)

df_2 \ df_1 (Denominator Degrees of Freedom \ Numerator Degrees of Freedom)	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799


المراجع

المراجع العربية

1. الاستاذ الدكتور محمد صبحي ابو صالح ، الطرق الاحصائية ، دار اليازوري العلمية للنشر والتوزيع ، الاردن ، 2009
2. السعدي ، سليم ذياب ، مبادئ علم الاحصاء ، دار الكتاب الجديد ، بيروت ، لبنان ، 2001
3. النعيمي ، محمد عبدالعال و الفضل مؤيد ، الاحصاء المتقدم في دعم القرار ، الوراق للنشر والتوزيع ، عمان ، الاردن ، 2007 .
4. الكيخيا ، نجاة رشيد ، اساسيات الاستنتاج الاحصائي ، دار المريخ للنشر ، المملكة العربية السعودية ، 2006 .
5. الصياد ، جلال مصطفى ، الاستدلال الاحصائي ، دار المريخ للنشر ، جدة ، المملكة العربية السعودية ، 1993 .
6. الصياد ، جلال مصطفى و حبيب محمد الدسوقي ، مقدمة في الطرق الاحصائية ، دار حافظ للنشر والتوزيع ، المملكة العربية السعودية .
7. عودة ، أحمد عودة ،الاحصاء الوصفي والاستدلالي ، دار الفلاح للنشر والتوزيع ، عمان الاردن ، 2014 .
8. كنجو ،أنيس أسماعيل واخرون ، مبادئ الاستدلال الاحصائي ، جامعة الملك سعود ، 2005
9. الشيحة عبدالله عبدالكريم ، أسس نظرية التقدير ، جامعة الملك سعود ، 2007
10. مصطفى ، عبدالحميد ، الاستدلال الاحصائي (1) نظرية التقدير ،مجموعة النيل العربية ، 2000

المراجع الاجنبية

1. Gupta S.P Statistical Methods ,Sultan Chand & Sons Publisher , New Delhi , India , 1994
2. Harry Frank and Steven C. [illegible]n , Statistics , Concepts and applications , Cambridge University Press , London , 1994

3. Gupta S.P , Practical Statistics , S. Chand and company LTD , New Delhi , 1998
4. Richard I .Levin & David S. Rubin , Statistics for Management , prentice ,Hall of India , New Delhi , 1992
5. Jane Miller , Statistics for Advance level , Cambridge University Press, Great Britain , 1994



Berlin10315 Gensingerstr :112
Tel :0049-code Germany
54884375-030
91499898-030
86450098-030
book@democratica.de